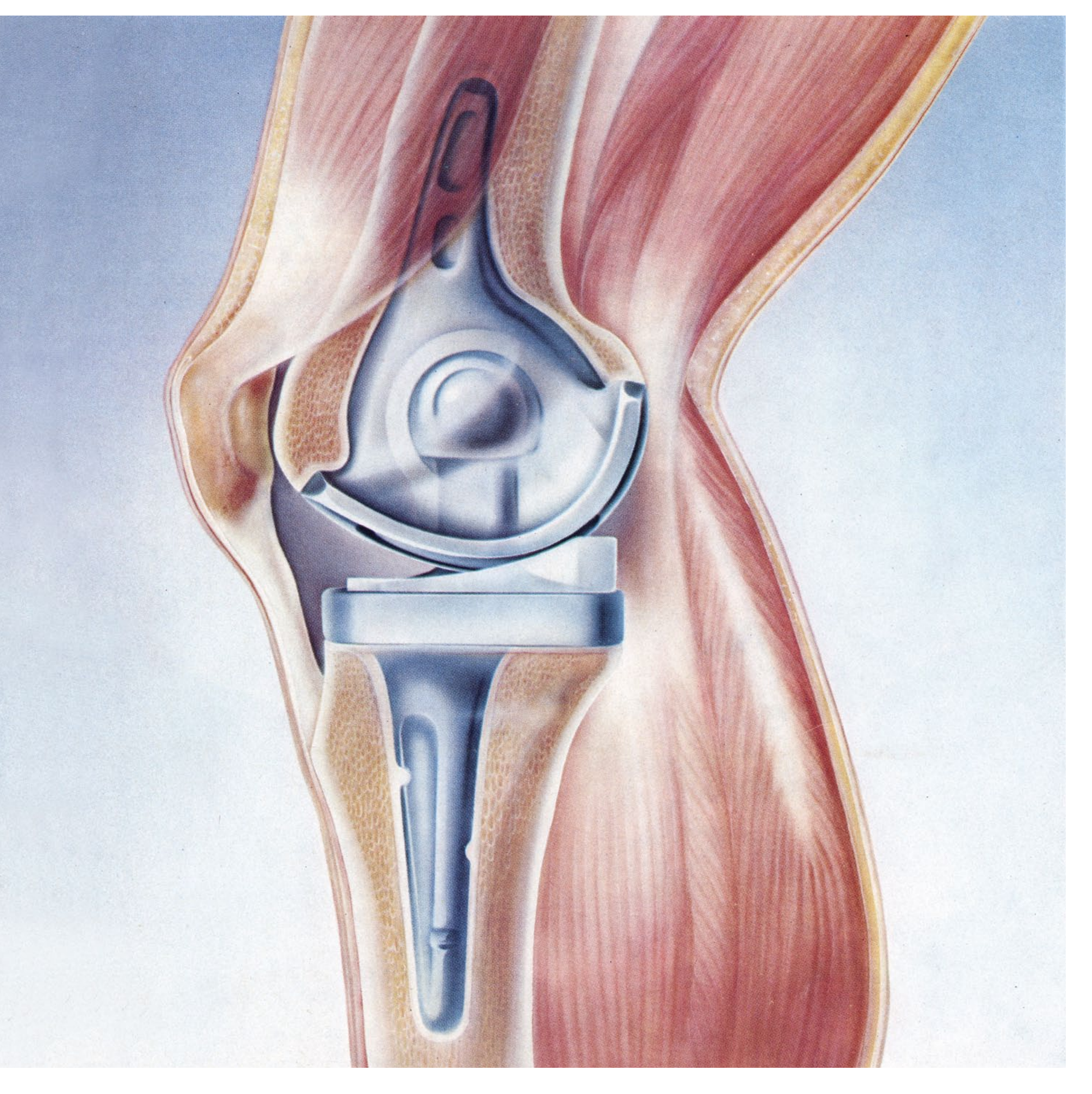
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



RODILLAS ARTIFICIALES

Marzo 1978

170 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 18
Marzo 1978

INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29
ESPAÑA

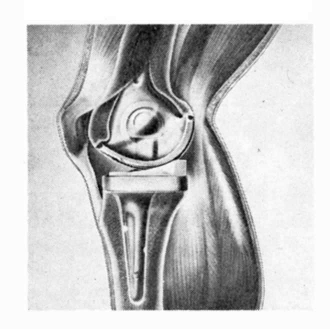
Colaboradores de este número:

Asesoramiento y traducción:

Enric Aulí: *La cuestión del dióxido de carbono;* Lino Torre: *La sustitución quirúrgica de la articulación de la rodilla;* Carles Oppenheimer: *La estructura tridimensional del ARN de transferencia;* José Javier Labay: *La estructura del medio interestelar;* José Gaspar Loren: *El mecanismo de adherencia en las bacterias;* Luis Bou: *La eficiencia de los algoritmos;* Antonio Blanco: *La Cartago romana;* Luis Bou: *Juegos matemáticos;* José E. Myro y Alicia Menéndez: *Taller y laboratorio;* Gerardo Espinosa: *Ciencia y sociedad.*

Libros:

Andrés Lewin y Andrés Sánchez Cascos.



LA PORTADA

La ilustración de la portada muestra una articulación de rodilla artificial, llamada rodilla esferocéntrica, creada en el Centro Médico de la Universidad de Michigan como prótesis para personas cuya artritis hubiese llegado a la etapa de grave invalidez (véase "La sustitución quirúrgica de la articulación de la rodilla", por David A. Sonstergard, Larry S. Matthews y Herbert Kaufer, en este número). El componente dispuesto en el fémur es básicamente una concavidad, en la cual se aloja una esfera que se alza desde el componente inserto en la tibia. En la parte inferior del componente femoral hay un par de carriles por los que se deslizan unos patines del componente tibial y que ayudan a controlar el movimiento de la rodilla. Los componentes van cementados en las partes óseas esponjosas con polimetil metacrilato.

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria, 235-239
Barcelona-29 (España)

Distribución para España:

Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:

Editorial Labor, S.A.
Calabria, 235-239
Barcelona-29 (España)

Precio de venta de este número:

España: 170 pesetas
Extranjero: 2 U.S. $

Condiciones de suscripción:

España:
Un año (12 números): 1870 pesetas
Extranjero:
Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
190 pesetas

Ejemplar atrasado extraordinario:
270 pesetas.

Difusión controlada por **OJD**

PROCEDENCIA DE LAS ILUSTRACIONES

Dibujo de la portada de Carol Donner

Página	Fuente	Página	Fuente
7	J. M. Valderas	60	Carl Heiles, Universidad de California en Berkeley, y Edward B. Jenkins, Princeton University
8-14	F. J. Ynduráin	61	Allen Beechel
17-18	National Aeronautics and Space Administration	62	© National Geographic Society-Palomar Observatory Sky Survey
19-25	Jerome Kuhl	63	Allen Beechel
29	University of Michigan Medical Center	66-67	J. W. Costerton Universidad de Calgary
30-31	Carol Donner	68	Ilil Arbel
32	James Young	69	J. W. Costerton Universidad de Calgary
33-34	Carol Donner	70	Ilil Arbel
35	James Young	71	J. W. Costerton Universidad de Calgary
43	Allen Beechel	72-73	Ilil Arbel
44	John Wiley & Sons, Inc. *(arriba)*. Allen Beechel *(abajo)*	74-75	J. W. Costerton Universidad de Calgary
46-47	Allen Beechel	79-90	Tom Prentiss
48-49	Sung Hou Kim, Duke University School of Medicine	92-94	Francis W. Kelsey Museum of Ancient and Mediaeval Archaeology, Universidad de Michigan
50-53	Allen Beechel	95-99	Lorelle M. Raboni
54-55	Carl Heiles, Universidad de California en Berkeley, y Edward B. Jenkins, Princeton University	104	Ilil Arbel
56-57	Allen Beechel	105-109	Ben Rose
58-59	Carl Heiles, Universidad de California en Berkeley, y Edward B. Jenkins, Princeton University *(arriba);* Allen Beechel *(abajo)*	110	Ilil Arbel
		112-116	Michael Goodman

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-29
Cayfosa. Santa Perpetua de Moguda.
Barcelona.

Printed in Spain - Impreso en España

Cartas

Señor Director:

El fascinante artículo de Norman H. Horowitz "La búsqueda de vida en Marte", (INVESTIGACION Y CIENCIA, enero 1978), incluye una breve exposición de la química del carbono, en la que se dice: "Entre los elementos, el carbono es único en el número, variedad y complejidad de compuestos que puede formar... [Ningún otro] tiene la capacidad del carbono para formar esas grandes y complejas estructuras que son tan estables."

Ahora se sabe que el boro, vecino inmediato del carbono, emula a éste y en ciertos aspectos hasta lo supera en su capacidad de formar redes moleculares estables de enlaces covalentes. Existen las familias de aniones borano tales como el $B_{12}H_{12}^{2-}$ (un icosaedro), carboranos [por ejemplo $(C_5H_5)\ CoC_2B_5H_7$], metaloboranos, fosfacarboranos y numerosas otras familias que abarcan miles de compuestos. En cuanto a estabilidad, el p-carborano ($1,12$-$C_2B_{10}H_{12}$) ¡soporta temperaturas de más de 700 grados Celsius y no es atacado por el ácido sulfúrico hirviente!

Las geometrías de los sistemas de jaula del boro son muy diferentes de los acostumbrados en la química de hidrocarburos, y no discuto el aserto principal del profesor Horowitz, a saber, que es improbable la existencia de formas biológicas no basadas en el carbono; en este respecto es preciso reconocer la singularidad del carbono mientras no se demuestre lo contrario. Pero desde un punto de vista estrictamente químico, hay ahora dos elementos –el boro y el carbono– que son capaces de formar armazones extensos y variados. De hecho, en sus compuestos en jaula, el boro exhibe un grado de versatilidad incluso mayor que el del carbono.

RUSELL N. GRIMES

Catedrático de Química
Universidad de Virginia

Señor Director:

Cualquier residente de Wisconsin que osara afirmar que el gato doméstico emula a su camarada bovino de granero en cuanto a importancia económica sería realmente insensato. Pero sostener que los gatos "nunca tuvieron ningún tipo de significado económico" ["Gatos y vías comerciales", por Neil B. Todd; INVESTIGACION Y CIENCIA, enero, 1978] es inverosímil en sentido contrario. Parece haber acuerdo de que en los graneros de viejas ciudades de Libia y regiones cercanas fue donde por primera vez se encerró a una, o posiblemente varias, especies de pequeños *Felis*. Allí se les encargaba de proteger las mercancías contra los hambrientos roedores durante los inviernos improductivos y los largos asedios que hubiese. Si evitar que toda una ciudad perezca de hambre no tiene significado económico, tal noción se torna realmente elusiva. Los gatos fueron más tarde trampas de ratones para egipcios y griegos, aunque al parecer no lo fueron para los romanos, y los historiadores mencionan su posible función de controladores de la peste bubónica.

JACK P. HAILMAN

Catedrático de Zoología
Universidad de Wisconsin

Señor Director:

En el artículo "Gatos y vías comerciales", de Neil B. Todd, al analizar la variedad tabí manchado del gato doméstico, se llega a la conclusión de que existe un foco inglés, y probablemente un origen inglés, para la diseminación de este mutante, que habría penetrado a través de Francia y luego rápidamente a través del Mediterráneo hacia Oriente. El profesor Todd observa asimismo que si bien no es claro el motivo de la ventaja selectiva del tabí manchado, la variedad se está diseminando como una epidemia que afectará a todas las poblaciones gatunas.

En la literatura se hallan ciertas pruebas de que el tabí quizá no sea originario de Inglaterra sino del foco oriental mencionado por el profesor Todd. De modo que quizá se haya difundido hacia el poniente y no hacia el levante.

John Aubrey, hidalgo inglés que vivió entre 1626 y 1697, fue anticuario, miembro de la Royal Society, y compiló y registró un vasto acervo de datos biográficos y de otras especies que, publicados con el título *Aubrey's Brief Lives*, ha demostrado ser valioso como fuente documental para historiadores interesados en la Inglaterra del siglo diecisiete. Según Aubrey, el gato tabí era novedad en Inglaterra a principios del s. XVII. Anotaba que a W. Laud, graduado de Cambridge y amante de los gatos, le regalaron unos gatos chipriotas, es decir gatos tabí, que en aquel entonces (1637) se vendían a cinco libras esterlinas cada uno. Esto, según parece, originó una nueva moda, y un cambio del que se quejaba Aubrey amargamente. El tabí suplantó rápidamente al "gato inglés común", del que Aubrey decía recordar muy bien que era blanco con algunas pintas azuladas. "Su raza o especie está casi desaparecida ahora", se lamentaba Aubrey.

GEOFFREY WOOD

Falls Church, Virginia

Señor Director:

El profesor Hailman adopta una perspectiva de la importanuia económica del gato doméstico de mayor amplitud que la que yo normalmente admitiría. Pero tengo mis dudas acerca de que los gatos hayan salvado alguna vez del hambre a una ciudad, exceptuando quizá el hecho de constituir una fuente alimentaria de reserva para la población humana. Así ocurrió en la isla de Rodas durante la ocupación alemana en la II Guerra Mundial, y probablemente también en Leningrado durante el asedio alemán. En estos casos, la función de los gatos no se diferenciaba mucho de la de una pulpa de papel para empapelar, o cortezas de árbol y cuero de calzado. En cuanto a controlar la diseminación de la peste negra, las estadísticas de mortalidad no parecen sugerir que el gato fuese notablemente eficaz.

El señor Wood introduce ciertas seductoras briznas de prueba, pero encuentro que son en buena medida imposibles de interpretar. El "gato inglés común" que era blanco con algunas pintas azuladas debe estar totalmente "desaparecido" ahora, porque no puedo decir que haya visto nunca uno de esos bichos. En cuando al tabí usurpador, ¿qué clase de tabí era? He descubierto que la mayoría de los dueños de gatos no hacen un verdadero distingo entre manchado y a rayas.

La información reciente acerca de las frecuencias del tabí manchado en Irán, Paquistán y la Unión Soviética indica que el foco se sitúa en Irán oriental (Mashhad) y que la zona y la intensidad de su aparición son mucho más restringidas de lo que son en Europa occidental. Eso sugiere un origen más reciente en Oriente, probablemente a causa de una mutación independiente, aunque Mashhad, a horcajadas sobre la ruta de la seda, puede haber recibido inicialmente algunos tabíes manchados.

NEIL B. TODD

Carnivore Genetics Research Center
Newtonville, Massachusetts

Los autores

FRANCISCO JOSE YNDURAIN ("Teorías unificadas de las interacciones fundamentales") es catedrático de Física de la Universidad Autónoma de Madrid desde 1971. Se licenció por la Universidad de Zaragoza en 1962, doctorándose en la misma en 1964, a los veinticuatro años de edad. Pertenece a numerosas sociedades científicas y ha enseñado en distintas instituciones europeas y americanas. Sus campos de interés se centran en la teoría general de colisiones a alta energía y, desde 1975, en las teorías gauge de partículas elementales.

GEORGE M. WOODWELL ("La cuestión del dióxido de carbono") dirige el Ecosystems Center en el Marine Biological Laboratory de Woods Hole, Massachusetts. Cursó estudios en el Dartmouth College y se doctoró en Botánica por la Duke University en 1958. Pasó tres años enseñando en la Universidad de Maine en Orono, que abandonó en 1961 para entrar a formar parte del equipo del Brookhaven National Laboratory. En enero de 1975 tomó posesión de su actual cargo. Las investigaciones del profesor Woodwell se han centrado en el estudio de la estructura y la función de las comunidades naturales, especialmente bosques y estuarios, y en los efectos sobre el medio ambiente de las radiaciones ionizantes y los pesticidas.

DAVID A. SONSTEGARD, LARRY S. MATTHEWS y HERBERT KAUFER ("La sustitución quirúrgica de la articulación de la rodilla humana") son cirujanos ortopédicos y bioingenieros de la Universidad de Michigan. Sonstegard es profesor asociado de Mecánica y Biomecánica aplicada a la cirugía. Se graduó en la Escuela de Minería y Tecnología de Dakota del Sur y consiguió el doctorado en Ingeniería mecánica por Michigan en 1965. Después de varios años dedicado a la investigación mecánica, volvió a Michigan en 1970, para dedicarse a estudios biomecánicos. Matthews es profesor asociado de Ortopedia. Estudio en Harvard, en donde se doctoró en Medicina en 1963 y regresó a Michigan para realizar su internado y residencia en Ortopedia. Después de trabajar durante dos años como cirujano ortopédico del Ejército, dedicó un año a trabajos de investigación biomecánica en el laboratorio de Carl Hirsch en el Instituto Carolino de Estocolmo, regresando a Michigan en 1971. Kaufer es cirujano y jefe de Clínica ortopédica. Estudió en Michigan, donde consiguió el doctorado en 1959. Concluido allí su período de internado y residencia en Ortopedia, sirvió en las Fuerzas Aéreas como jefe de cirugía ortopédica. En 1966 se incorporó a la Universidad de Michigan

ALEXANDER RICH y SUNG HOU KIM ("La estructura tridimensional del ARN de transferencia") son biofísicos con un interés especial por el mecanismo de la síntesis proteica. Rich enseña Biofísica en la Cátedra Thomson Sedgwick del Instituto de Tecnología de Massachusetts. Cursó estudios en Harvard College y en la Harvard Medical School, donde se doctoró en Medicina en 1949. Durante cinco años trabajó como postdoctorado en investigaciones sobre químico-física con Linus Pauling en el Instituto de Tecnología de California. Desde 1958 pertenece al claustro docente del M.I.T. Kim es profesor asociado de Bioquímica en la Duke University School of Medicine. Natural de Corea, realizó sus trabajos de graduado en la Universidad Nacional de Seúl y obtuvo el grado de doctor en Cristalografía por la Universidad de Pittsburgh en 1966. Durante los seis años siguientes realizó su labor postdoctoral en el laboratorio de Rich, del M.I.T. En 1972 se trasladó a Duke.

CARL HEILES ("La estructura del medio interestelar") es profesor de Astronomía en la Universidad de California en Berkeley. Realizó sus estudios de postgraduado en la Cornell University, doctorándose en Astronomía por la Universidad de Princeton en 1966, año en que se incorporó a la Facultad de Berkeley.

J.W. COSTERTON, G. G. GEESEY y K.-J. CHENG ("El mecanismo de adherencia en las bacterias") son microbiólogos cuyo campo de investigación es la pared celular bacteriana. Costerton es profesor de Biología en la Universidad de Calgary en Canadá. Realizó sus estudios de postgraduado en la Universidad de la Columbia Británica y obtuvo su doctorado en Bacteriología por la Universidad de Western Ontario en 1960. Geesey es profesor contratado del departamento de Bioquímica y Microbiología en la Universidad de Victoria en la Columbia Británica. En 1975 obtuvo su doctorado en Microbiología marina por la Universidad estatal de Oregón. A partir de este momento, trabaja como investigador postdoctoral asociado en el laboratorio de Costerton. Cheng es un especialista en la microbiología del rumen en el Departamento de Agricultura Canadiense. Nacido en Taiwan, se formó en la Universidad Nacional de Taiwan y en la Universidad de Saskatchewan, donde obtuvo su doctorado en Microbiología en 1969. Se incorporó al laboratorio de Costerton en la Universidad de McGill en el año 1968, desde donde pasó a ocupar el puesto actual.

HARRY R. LEWIS y CHRISTOS H. PAPADIMITRIOU ("La eficiencia de los algoritmos") son profesores ayudantes de Ciencias de la computación en el Center for Research in Computing Technology de la Universidad de Harvard. Lewis fue alumno del Harvard College, donde estudió matemática aplicada, interesándose después en análisis de grafos mediante ordenador. Tras graduarse "summa cum laude" en 1968 prestó servicios en el Public Health Service, diseñando sistemas y programas de aplicación para un gran computador perteneciente a los National Institutes of Health. Volvió a la universidad para continuar estudios de postgraduado en matemática aplicada, recibiéndose de doctor en 1974. Papadimitriou, nacido en Atenas, realizó sus primeros estudios universitarios en ingeniería eléctrica en la Universidad Politécnica de su país. Se recibió de doctor por la Universidad de Princeton en 1976, pasando a formar parte del claustro de Harvard ese mismo año.

JOHN H. HUMPHREY y JOHN GRIFFITHS PEDLEY ("La Cartago romana") son arqueólogos clásicos de la Universidad de Michigan. Humphrey es profesor ayudante del departamento de estudios clásicos y conservador ayudante del Museo Francis W. Kelsey de Arqueología antigua y medieval, de la Universidad. Estudió arqueología clásica en la Universidad de Cambridge, donde recibió su grado con honores máximos condecorados. Se doctoró en Filosofía en Bryn Mawr College con una tesis sobre los circos romanos. Pedley es profesor de Arqueología clásica y de Griego y también director del Museo Kelsey. Nacido en Inglaterra, obtuvo el grado de licenciado en Cambridge en 1959 y pasó los cuatro años siguientes con becas en la Universidad de Harvard y en la Escuela Americana de Estudios Clásicos en Atenas. Volvió a Harvard y después se incorporó a la Facultad de Michigan. Pedley ha participado en excavaciones en Sardes, Apollonia y Cartago.

Teorías unificadas de las interacciones fundamentales

Los recientes descubrimientos experimentales en física de partículas elementales sugieren y confirman teorías unificadas de sus interacciones y de la estructura última de la materia

Francisco José Ynduráin

Los últimos cuatro o cinco años han sido testigos de un fantástico florecimiento de trabajos, tanto experimentales como (sobre todo) teóricos, sobre los esquemas unificados de las interacciones fundamentales. Literalmente, miles de artículos han aparecido en revistas científicas sobre consideraciones teóricas y verificaciones experimentales de modelos en los que dos o más de las fuerzas básicas que actúan sobre los últimos constituyentes de la materia (las partículas elementales) se tratan no como entes distintos, sino como manifestaciones de un principio más fundamental que los subyace.

Como todo desarrollo científico, también éste tiene un origen doble: la aparición de esquemas teóricos coherentes, de los que se pueden extraer consecuencias verificables experimentalmente, y, sobre todo, la constatación de que los experimentos realizados confirman las predicciones teóricas. En nuestro caso, los hallazgos que están en la base de todo el movimiento son la detección, por un grupo de científicos del CERN, de la existencia de corrientes neutras a principios de 1973, y el descubrimiento por los grupos dirigidos por Richter y Ting de las partículas J o ψ, y, posteriormente, de toda una serie de las llamadas "partículas encantadas". Estos descubrimientos experimentales, predichos por teorías unificadas, son los que hacen posible que la búsqueda de dichos esquemas sea una tarea científica, y no mera especulación filosófica, por primera vez desde hace muchos años; y aunque no haya, hoy por hoy, un esquema teórico único, todos los modelos aceptados tienen en común la característica de la unificación.

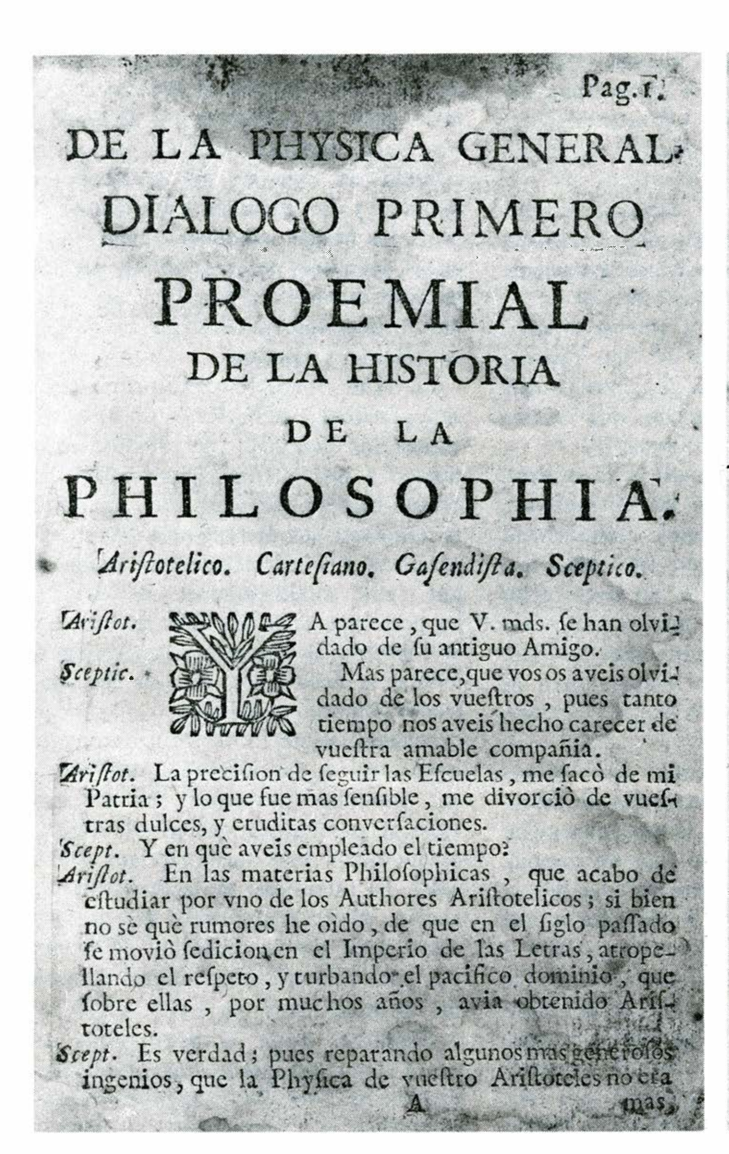
Estas características de fluidez y efervescencia son las que me han llevado a escoger el enfoque que utilizo en este artículo: en vez de presentar una discusión detallada de las teorías y resultados experimentales actuales, he preferido ofrecer una visión más histórica, encuadrando las ideas actuales como un episodio más de la busca permanente de unas estructuras elementales que "unifique" la aparente diversidad del mundo fenomenológico.

Perspectiva histórica

Como reconoce B. Russell en su libro *History of Western Philosophy* (Historia de la filosofía occidental), "Casi todas las hipótesis que han dominado la filosofía (que en el contexto incluye a la filosofía natural, i.e., la ciencia) moderna fueron imaginados por los griegos; su inventiva e imaginación en temas científicos apenas pueden exagerarse". Aunque sin caer en el error de pretender encontrar en la filosofía antigua más de lo que realmente contiene, no cabe duda de que vale la pena comenzar cualquier exposición histórica de un tema científico por las opiniones al respecto de los filósofos presocráticos; aunque sólo sea por darse cuenta de la cantidad de tiempo que la humanidad pensante lleva intentando dar respuesta a unos pocos problemas básicos, entre los que se cuenta el que nos ocupa.

Toda teoría unificada –toda unificación– tiene dos vertientes complementarias. En primer lugar, hay que reducir substancias que aparecen con propiedades diversas a agregados o diferentes estados de una materia primordial; y, en segundo lugar, es necesario explicar cómo las distintas fuerzas que actúan en la naturaleza son tan sólo distintas manifestaciones de una única fuerza (interacción, con la terminología moderna) básica. La manera de pensar, más geométrica que dinámica, de los griegos hace que sus disquisiciones se refieran, sobre todo, a la primera vertiente. Así, el primer filósofo griego del que tenemos noticia (Thales de Mileto, *circa* 585 a. de C.) aparte de sus importantes contribuciones a la geometría, ya propuso la hipótesis de que, en última instancia, "todo está hecho de agua". Aunque obviamente falsa, esta hipótesis no es tan descabellada como pueda parecer a simple vista. En efecto, el agua es el único elemento al que, con la primitiva tecnología de la época, podría hacérsele adoptar los tres estados de sólido, líquido y gas. Una tal teoría, por otra parte, no era difícil que se le ocurriera a un habitante de la Magna Grecia, conjunto de islas, penínsulas y puertos desparramados por el Mediterráneo; lo que ilustra bien la interdependencia de las condiciones sociales y geográficas y las hipótesis científicas.

El punto de vista de Thales puede contrastarse con el de Anaximandro (540 a. de C.), que mantuvo que la substancia primera no es ninguno de los elementos conocidos, sino que es una substancia que se encuentra oculta en todos ellos (compárese esta aserción con la moderna teoría de quarks). Esta interesante vía fue, sin embargo, abandonada por los filósofos posteriores que, con la excepción de los atomistas, consideraron que la substancia primera era una de las conocidas; y así, a través de Anaxímenes, Jenófanes y Heráclito, que propusieron el aire, la tierra y el fuego, respectivamente, llegamos a Empédocles que sintetizó todas estas hipótesis, postulando el cuarteto clásico: aire, agua, tierra y fuego. Este cuarteto fue aceptado por Aristóteles y, debido en parte a su autoridad, perduró por largos siglos (en particular a través de los alquimistas) hasta la aparición de la química en la edad moderna. A pesar del éxito de esta hipótesis, no es ésta lo más interesante de


Pag. 1.

DE LA PHYSICA GENERAL.

DIALOGO PRIMERO

PROEMIAL

DE LA HISTORIA

DE LA

PHILOSOPHIA.

Ariſtotelico. Carteſiano. Gaſendiſta. Sceptico.

Ariſtot. YA parece, que V. mds. ſe han olvidado de ſu antiguo Amigo.

Sceptic. Mas parece, que vos os aveis olvidado de los vueſtros, pues tanto tiempo nos aveis hecho carecer de vueſtra amable compañia.

Ariſtot. La preciſion de ſeguir las Eſcuelas, me ſacò de mi Patria; y lo que fue mas ſenſible, me divorciò de vueſtras dulces, y eruditas converſaciones.

Scept. Y en què aveis empleado el tiempo?

Ariſtot. En las materias Philoſophicas, que acabo de eſtudiar por vno de los Authores Ariſtotelicos; si bien no sè què rumores he oìdo, de que en el ſiglo paſſado ſe moviò ſedicion en el Imperio de las Letras, atropellando el reſpeto, y turbando el pacifico dominio, que ſobre ellas, por muchos años, avia obtenido Ariſtoteles.

Scept. Es verdad; pues reparando algunos mas generoſos ingenios, que la Phyſica de vueſtro Ariſtoteles no era mas,

A

INDICE

DE LOS DIALOGOS.

FIN.

HASTA EL SIGLO XIX no se dispuso de modelos científicos que explicaran los agregados constitutivos de la materia y las distintas fuerzas que actúan en la naturaleza. Desde Empédocles, quien redujo la substancia primera, *arjé*, a los cuatro elementos (agua, aire, tierra y fuego), se vinieron repitiendo los mismos principios y ejemplos, como se aprecia en la primera página e índice de la obra de Martín Martínez, publicada en 1730: *Filosofía scéptica, extracto de la física antigua y moderna, recopilada en diálogos entre un aristotélico, cartesiano, gasendista y scéptico.*

la física de Empédocles, sino el reconocer que no basta con identificar unas substancias primordiales, sino que hay que atender a lo que hemos llamado la segunda vertiente de la unificación: es decir, hay que explicar por qué estas substancias se combinan como lo hacen. Empédocles introdujo el rudimento del moderno concepto de fuerza o interacción, que él, de forma antropomórfica, identificó con "amor" y "odio": conceptos que básicamente están en el origen de la idea moderna de fuerzas "atractivas" y "repulsivas". El esquema de Empédocles es, por lo menos en la filosofía occidental, el primer esquema unificado coherente en el que cuatro elementos, y dos interacciones básicas, pretenden explicar la variedad del universo.

En realidad, todos estos modelos, por ingeniosos que resulten, son acientíficos: ni se basan en la experiencia ni se indican modos experimentales de verificar sus predicciones; para llegar a modelos científicos hay que esperar hasta el siglo XIX. A principios de este siglo se había conseguido formular la hipótesis atómica de forma científica, con Dalton (1808), aunque la confirmación experimental definitiva no fue obtenida sino mucho más tarde (J.J. Thomson 1897). Se conocían ya las leyes de enlace químico, quedando reducida la materia a los 92 elementos químicos fundamentales y las interacciones entre ellos a cuatro fuerzas básicas: eléctrica, magnética, gravitatoria y "de cohesión". (La fuerza nuclear fue descubierta mucho más tarde, esencialmente por Rutherford a principios del actual siglo). Estas cuatro fuerzas fueron objeto de unas unificaciones parciales durante el siglo XIX: unificaciones que contienen aspectos que perviven hasta en los modelos actuales.

La primera unificación es la de las fuerzas eléctricas y magnéticas debida a Maxwell; éste demostró que ambas son dos manifestaciones de un potencial básico: los fenómenos electromagnéticos se reducen a una propiedad intrínseca de la materia (carga eléctrica), y a las fuerzas creadas por estas cargas, en reposo (campos eléctricos) o en movimiento (campos eléctricos y magnéticos). Posteriormente, Millikan encontró que todas las cargas eléctricas son múltiplos enteros (positivos o negativos) de una carga fundamental que puede identificarse con la del electrón, y se denota por $-e$, o con la del positrón, y se denota por e. La fuerza que una carga elemental ejerce sobre otra (en reposo) y en el vacío viene dada por la conocida ley de Coulomb: $e^2/4\pi r^2$, donde r es la distancia entre las cargas. (Hemos utilizado el sistema de unidades usual en física subnuclear, en el que la velocidad de la luz, c, y la constante de Planck, $\hbar$, se toman como uni-

dad). Como se ve, la interacción entre dos cargas elementales está caracterizada por la constante de interacción (o *constante de acoplamiento,* nombre más usual): $\alpha = e^2/4\pi \simeq 1/137.036$, donde α es la constante en cuestión.

Esta intensidad es pequeña, pero resulta todavía muy grande si la comparamos con la gravitatoria; la atracción gravitatoria entre, por ejemplo, dos protones es: $G_N\ m_p^2/4\pi r^2$, siendo $G_N\ m_p^2 \simeq 10^{-39}$, es decir, unas 10^{36} veces menor. G_N es la constante gravitatoria de Newton y m_p la masa del protón.

La segunda unificación condujo a explicar las "fuerzas de cohesión" como residuo de las fuerzas electromagnéticas (fuerzas de Van der Waals). Para mostrar el mecanismo veamos un ejemplo sencillo (fuerza dipolar) sobre el que está calcado en buena parte el moderno modelo de quarks. Consideremos una molécula de hidrógeno. Esta molécula consta de dos átomos de dicho elemento, H' y H'' *(véase la ilustración de la página 9),* cada uno de los cuales está a su vez formado por un protón (símbolo p) y un electrón (símbolo e). Como las cargas de éstos son opuestas, se atraen y, al ligarse, forman los átomos de hidrógeno. Sin embargo, cada átomo es eléctricamente neutro: ¿cómo interaccionan para formar una molécula? La respuesta se ve claramente en la ilustración citada. El protón del segundo átomo, por ejemplo, está solicitado por dos fuerzas (aparte de la de su electrón acompañante que no influye en el argumento y por tanto no consideramos): la repulsión del protón del otro átomo, de magnitud $e^2/4\pi(R+r)^2$, y la de atracción del electrón del otro átomo, que vale $-e^2/4\pi R^2$ donde r es la distancia intraatómica entre protón y electrón y R es la distancia interatómica entre el electrón del átomo de H' y el protón del átomo de H''. Como vemos, no se cancelan exactamente, sino que queda una resultante, atractiva, que viene dada por la siguiente razón: $-e^2\ (r^2 + 2rR)/4\pi R^2\ (R + r)^2$.

Esta fuerza residual (fuerza dipolar, caso particular de las del tipo Van der Waals) explica la atracción entre los dos átomos H' y H'', atracción que los liga en la molécula de hidrógeno, H_2.

Es instructivo razonar a la inversa. A gran distancia *(R)* las interacciones entre átomos son muy complicadas, según acabamos de ver; ello indica que son objetos compuestos. En efecto, si los miramos a distancias más pequeñas encontramos sus constituyentes (electrones y protones) que interaccionan de acuerdo con la sencilla ley de Coulomb: $(e^2/4\pi r^2)$.

Como es lógico, todos estos esquemas fueron propuestos antes de descubrirse la relatividad y la mecánica cuántica; su adaptación a éstas no cambió, sin embargo, nada esencial. De hecho, lo que hizo estrecho el esquema en que dos interacciones (electromagnética y gravitatoria) explicasen toda la física fue el descubrimiento, al filo de la transición al siglo XX, de las nuevas interacciones (llamadas, con falta de imaginación, "débiles" y "fuertes" por ser, respectivamente, más débiles y fuertes que la interacción electromagnética), y no la adaptación a los requisitos generales relativistas o cuánticos. La adaptación del electromagnetismo, por ejemplo, es tan excelente que la teoría cuántica relativista (Quantum Electro-Dynamics, QED en inglés) es una de las teorías más perfectas que se conocen; estando de acuerdo las predicciones teóricas y las medidas experimentales hasta los límites conocidos (unas nueve cifras decimales exactas). Similarmente, la teoría cuántica no hizo sino perfeccionar el conocimiento de las fuerzas químicas (Heitler y London).

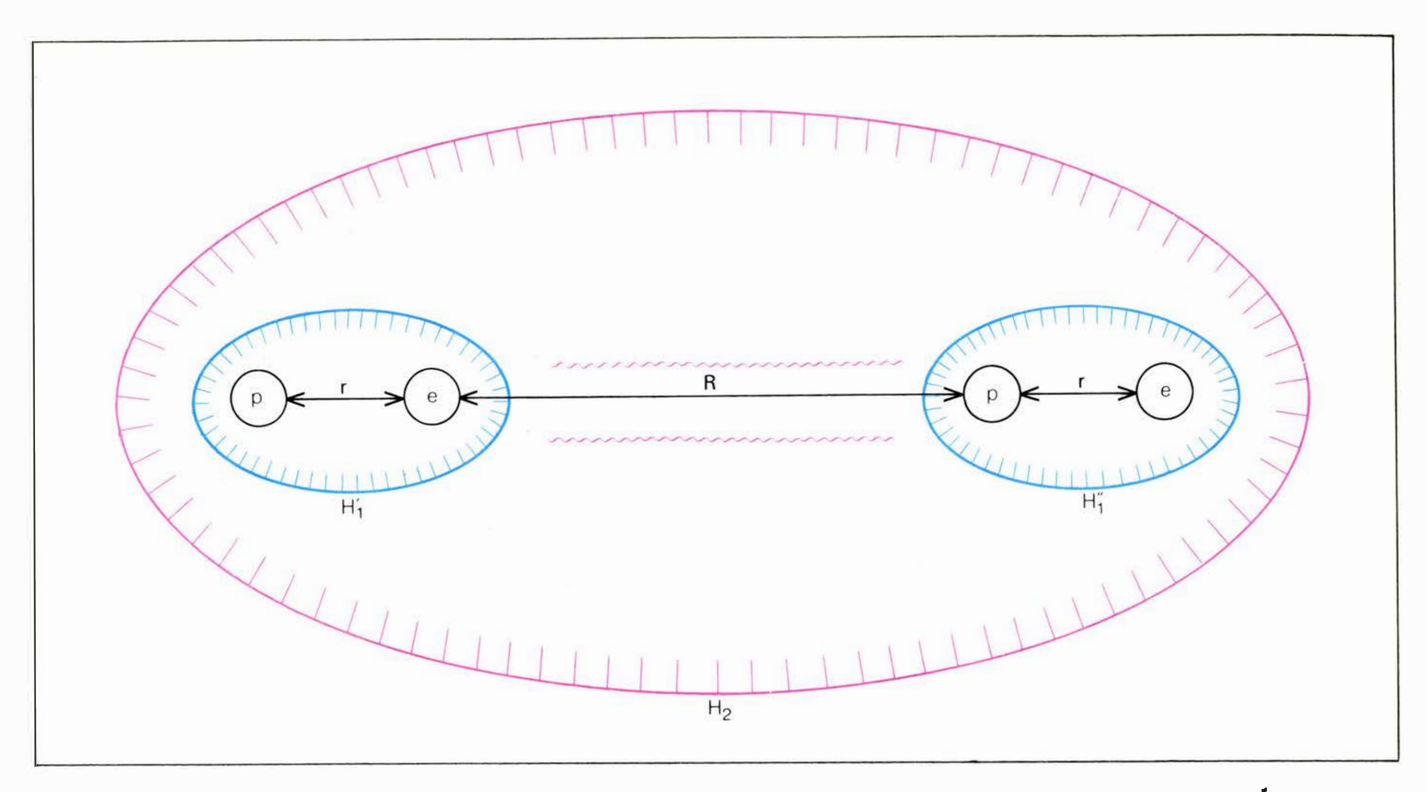
En 1896 Becquerel descubrió las desintegraciones radiactivas (desintegraciones β), fenómeno que desafió largo tiempo a una interpretación teórica. El primer modelo propuesto que tuvo éxito al comparar sus predicciones con la experiencia fue el iniciado por Fermi, y que, con modificaciones de detalle, ha pervivido hasta los años sesenta. En este modelo, se supone que la desintegración β se debe a que uno de los neutrones (símbolo n) del núcleo atómico se desintegra según el esquema siguiente: $n \rightarrow p + e + \nu$; ν es el símbolo del neutrino, partícula postulada por motivos teóricos por Pauli y encontrada experimentalmente más tarde (Reines y Cowan). La intensidad de la fuerza (interacción débil) responsable del proceso viene dada por la constante de Fermi, G_F, que, tomando como unidad la masa del protón, tiene un valor de $G_F = 0.0000102\ (m_p^{-2})$: es decir, unas mil veces más débil que las fuerzas electromagnéticas. Uno podría sentirse tentado de explicar esta fuerza como un residuo a la Van der Waals de las electromagnéticas: esto es imposible debido a que su alcance es pequeñísimo, compatible con cero dentro de los límites experimentales actuales y, desde luego, mucho menor que el tamaño del núcleo.

La *interacción fuerte* tiene su origen experimental en el descubrimiento de Rutherford de que los átomos contienen a todos los electrones en la periferia, y que todos los protones se encuentran en un núcleo de tamaño mucho más reducido que el del átomo. La carga de este núcleo es netamente positiva; para explicar por qué no se desintegra bajo el efecto de la repulsión eléctrica entre los protones que contiene, hace falta postular una nueva interacción (llamada por su gran intensidad *interacción fuerte*) que sólo afecta a los constituyentes del núcleo (pero no a los electrones), caracterizada por la constante de Yukawa, g_Y, de valor $g_Y^2/4\pi \simeq 14$. Compárese con la constante de acoplamiento electromagnético (α) o con la constante de Fermi (G_F), y se apreciará su enorme superioridad en punto a intensidad.

Para acabar este capítulo, que nos lleva a los albores de lo que podríamos

	NOMBRE	SIMBOLO	CARGA (*a*)	MASA (*b*)	INTERACCIONES (*c*)	DESINTEGRACION (*d*)
	FOTON	γ	0	0	e.m.	ESTABLE
LEPTONES	ELECTRON	e	−1	0,51	e.m.w.	ESTABLE
LEPTONES	MUON	μ	−1	106	e.m.w.	$e+\bar{\nu}_e+\nu_\mu$
LEPTONES	NEUTRINOS {ELECTRONICO, MUONICO}	ν_e ν_μ	0	0(*e*)	w	ESTABLES (*f*)
HADRONES	PROTON	p	+1	938,3	s,e.m.,w	ESTABLE (*f*)
HADRONES	NEUTRON	n	0	939,6	s,w	$p+e+\bar{\nu}_e$
HADRONES	PIONES {NEGATIVO NEUTRO POSITIVO}	π^- π^0 π^+	−1 0 +1	139,6 135,0 139,6	s,e.m.,w s,w s,e.m.,w	$\mu+\bar{\nu}_\mu$ $\gamma+\gamma$ $\mu+\bar{\nu}_\mu$
HADRONES	ETA	η	0	548,8	s,w	$\gamma\gamma, 3\pi$

PARTICULAS "NORMALES". A cada partícula (de cualquier tipo) le corresponde una antipartícula, que denotamos poniendo una barra sobre el símbolo correspondiente. Partícula y antipartícula coinciden a veces: $\gamma=\gamma$; $\pi^0=\pi^0$. Respecto a las letras: *a)* Unidad de carga: la del protón. *(b)* Unidad de masa: 1MeV. *(c)* e.m.: electromagnéticas; w: débiles; *s:* fuertes. Sólo se explicitan las interacciones directas, no las inducidas. *(d)* Sólo se cita la forma de desintegración más común. *(e)* En algunas teorías, los neutrinos tienen una pequeña masa. Experimentalmente, $m_{\nu e} < 0{,}00006$ MeV, $m_{\nu\mu} < 0{,}65$ MeV. *(f)* En algunas teorías, los neutrinos y/o el protón son inestables aunque con enormes vidas medias. Experimentalmente, la vida media del protón es como mínimo de 10^{20} años. Por ahora no se conoce límite experimental útil a la vida de los ν (neutrinos).



MODELO de la molécula de hidrógeno, H² *(en rojo).* Los "residuos" de las fuerzas electromagnéticas elementales entre los elementos *(p y e)* que componen los átomos H₁' y H₁" *(en azul)* ligan a éstos. La interacción entre H' y H" es un residuo de las interacciones más básicas, de sus constituyentes.

llamar la problemática actual, es conveniente hacer una digresión sobre la naturaleza de las distintas fuerzas discutidas y una compilación de los constituyentes elementales ya encontrados. Con respecto a lo primero, la noción de fuerza se asocia con la de carga y campo (que es una generalización de la noción de campo electromagnético). La relatividad, sin embargo, al no admitir la acción a distancia, hace que el campo tenga una existencia propia; la mecánica cuántica, que implica un dualismo entre onda (y campo) y partícula, dice que cada campo tiene asociada una partícula. Asimismo, la relatividad y la mecánica cuántica ligan el alcance de una fuerza (campo) con la masa de la partícula asociada. Supongamos que una partícula A crea un campo, asociado con la partícula P *(véase la ilustración de la página 12).* Para que el campo actúe hace falta que se cree la partícula P. La energía mínima necesaria para esto es, según el principio de equivalencia de Einstein, la propia masa m_P de la partícula P. Esta creación "espontánea" es posible porque la energía no está bien definida, sino que su fluctuación, ΔE, está ligada con la distancia d sobre la que está extendida por la famosa relación de incertidumbre de Heisenberg, $\Delta E \cdot d \simeq \hbar c$.

Por tanto, como la fluctuación mínima para crear la partícula P es m_P, el campo que ésta transporta sólo será efectivo a la distancia d (alcance de la interacción), dado por la expresión $d = 1/m_P$ (utilizando el sistema de unidades con $\hbar = c = 1$). El campo electromagnético tiene alcance infinito; luego su partícula asociada (el fotón, designado con el símbolo γ) ha de poseer masa cero. Efectivamente, las mediciones experimentales más precisas dicen que m_γ es menor que 10^{-19} veces la masa del electrón. El alcance de las interacciones débiles es extremadamente pequeño. De hecho, no ha podido ser medido; sólo se conoce un límite para la masa de las partículas asociadas (partículas W): m_W mayor que unas treinta veces la masa del protón. Finalmente, las interacciones fuertes tienen como partículas asociadas los llamados mesones (en particular el pion, símbolo π) de masa unas doscientas ochenta veces la del electrón, lo que da un alcance de unos 10^{-13} centímetros. Compárese con el tamaño de un átomo, unos 10^{-8} centímetros, cien mil veces mayor que el núcleo.

Partículas extrañas, simetrías superiores y quarks

La fauna de las partículas conocidas antes de la segunda guerra mundial está resumida en la ilustración de la página 8; nótese que a los ya discutidos hemos añadido el muon (símbolo μ) de propiedades exactamente análogas a las del electrón, pero de masa 200 veces mayor y la partícula η (de propiedades análogas a las del pion; ver más adelante); y hemos distinguido dos neutrinos, ν_e y ν_μ: el primero acompaña siempre al electrón, el segundo al muon. Una primera agrupación distingue entre leptones, que son los más ligeros y sólo tienen interacciones electromagnéticas y débiles y hadrones (los más pesados) que, además, presentan interacciones fuertes. Aparte queda el fotón.

Con las partículas e interacciones resumidas en la ilustración de la página 8 se explican, al menos en principio, todos los fenómenos del mundo que nos rodea. En este sentido, tres interacciones (más la gravitatoria, que a propósito dejamos para el final del artículo) y unos diez constituyentes elementales no parece demasiada complicación (compárese por ejemplo con los cuatro elementos y las dos fuerzas de Empédocles) y, aunque incompleta, la unificación de la descripción de la materia es verdaderamente impresionante. Sin embargo, un vistazo a la tabla de la ilustración recién citada nos indica que deberá ser posible simplificar más. Por ejemplo, el neutrón y el protón tienen propiedades extremadamente análogas (excepto la carga eléctrica); y lo mismo pasa con los leptones. Los intentos de unificar estas partículas son bastante antiguos (años 30, 40 y 50); en el caso de los leptones chocan siempre con dificultades que sólo se han superado recientemente, y lo discutiremos más adelante; en el caso de los nucleones (protón y neutrón) y piones esto llevó (Condon y Shortley, Heisenberg) al llamado esquema de *isospín* que describiré en algún

detalle porque su generalización forma la base de todas las unificaciones actuales. (Isospín es una voz compuesta por contracción de los términos "isotópico", que ocupa el mismo rango, es decir, que tiene las mismas propiedades, y "spin", giro en inglés).

La hipótesis de Heisenberg consistía en suponer que, además del espacio ordinario, existe un espacio interno a las partículas elementales. En el caso del isospín, este espacio es de dos dimensiones (complejas); además, se supone que las interacciones fuertes no son capaces de distinguir entre las distintas direcciones en este espacio interno. Por tanto, todas las propiedades de tales interacciones serán invariantes bajo cambios de dirección en el espacio interno. Estos cambios forman lo que los matemáticos llaman una estructura de grupo que, si el espacio tiene dos dimensiones (complejas) es el llamado grupo SU(2). [En el caso del espacio ordinario, con tres dimensiones reales, el grupo correspondiente es el de rotaciones ordinarias, SO(3) en nomenclatura matemática.] La manifestación física de este espacio se hace por lo que se denomina representaciones del grupo, concepto que no es fácil de explicar en términos no matemáticos. El resultado es que cada grupo da lugar a una serie de representaciones; a cada representación le corresponde una partícula elemental, que puede manifestarse en tantos estados distintos como dimensiones tenga la representación. En el caso del grupo SU(2), tenemos representaciones con dimensión 1, 2, 3, 4... Una partícula correspondiente a la representación de dimensión 1 es la llamada *mesón eta* (símbolo η), que no se conocía aún al ser propuesto el modelo; la de dimensión 2 al nucleón, en sus dos posibles estados (protón y neutrón), la de dimensión 3 al pion (en sus tres posibles estados). Las partículas correspondientes a la representación de dimensión 4 se descubrieron posteriormente (resonancias Δ), etcétera.

Está claro que el modelo no puede ser exacto. En efecto, el protón y neutrón pueden ser iguales, con respecto a las interacciones fuertes, pero como uno tiene carga eléctrica y el otro no, resulta que las interacciones electromagnéticas los distinguen. Aunque el efecto es pequeño (la interacción electromagnética es, como hemos visto, unas mil veces más débil que la fuerte) no cabe duda que nos encontramos con un modelo que es sólo aproximado: a este fenómeno se le llama de ruptura de simetría. Esto presentaba dificultades conceptuales difíciles de resolver. Pero es que, además, el esquema fue sobrepasado por descubrimientos totalmente inesperados, que pasamos a discutir.

Para empezar, tenemos las partículas extrañas. Primero en los rayos cósmicos y luego en cuanto se construyeron aceleradores con energía suficiente en el laboratorio, se observó la existencia de partículas (llamadas, con una evidente falta de imaginación, "partículas extrañas") con propiedades muy peculiares. Cuando estas partículas interaccionan entre sí o con las ordinarias, lo hacen con una fuerza similar a la de Yukawa, pero cuando se desintegran en partícuias "ordinarias" en procesos tales como $\Lambda \to p + \pi^-$, $K^- \to \mu + \bar{\nu}_\mu$, la fuerza de la interacción correspondiente varía entre G_F y unas veinticinco veces menor que G_F. (Algunas de las propiedades de las partículas extrañas pueden verse en la ilustración de esta página.) ¿Cuál es la solución a este rompecabezas? Influidos quizá por la teoría del isospín, Pais, Nambu, Nijishima y Gell-Mann postularon que, además de otras propiedades comunes con las partículas ordinarias, las partículas extrañas tienen una propiedad intrínseca (hipercarga o extrañeza) que se conserva en interacciones fuertes o electromagnéticas, si bien las interacciones débiles pueden violar. De esta manera, como por hipótesis todas las partículas ordinarias *(véase la ilustración de la página 8)* tienen extrañeza cero, procesos tales como colisión elástica *K-p* vienen afectados por la interacción fuerte, pero desintegraciones como $\Lambda \to p + \pi^-$, $K^- \to \mu + \bar{\nu}_\mu$, en que la extrañeza cambia, tienen necesariamente que ser débiles.

Aún quedaba sin embargo un problema. Las desintegraciones débiles de las partículas extrañas son de dos tipos: aquellas en las que la extrañeza se conserva, tales como $\Sigma^- \to \Sigma^0 + e + \bar{\nu}e$, y aquellas en las que cambia, como las $\Lambda \to p + \pi^-$, $K^- \to \mu + \bar{\nu}_\mu$. Los tipos de desintegraciones en las que la extrañeza (S) se conserva tienen una intensidad $G(\Delta S = 0) \simeq G_F$, mientras que las que cambian la extrañeza viene caracterizada por $G(\Delta S \neq 0) = (0.2) \times G_F$. La unificación de ambos tipos de procesos fue dada por Cabibbo, quien supuso la existencia de una única fuerza débil elemental, G_F, y, que esta fuerza se reparte entre procesos con cambio de extrañeza, y procesos sin cambios de extrañeza: de esta forma, $G(\Delta S = 0) = G_F \cos\theta$, $G(\Delta S \neq 0) = G_F \operatorname{sen}\theta$; donde θ (experimentalmente, $\theta = 0{,}2$ radianes) es el "ángulo de reparto" que indica qué proporción de fuerza débil corresponde a cada tipo de proceso. (La manera más clara de ver esto es dentro del modelo quark.) Esta hipótesis ha sido verificada con gran precisión con los cientos de desintegraciones conocidos actualmente, en el sentido de que todas las desintegraciones observadas experimentalmente se explican con el esquema; pero presenta un problema, irrelevante en apariencia, que resulta ser clave.

El problema es el llamado de las corrientes neutras con cambios de extrañeza. Todos los procesos de desintegración de partículas extrañas observados experimentalmente se caracterizan por presentar, además de un cambio eventual de extrañeza, una transferencia de carga eléctrica, del hadrón inicial a los leptones finales. Por ejemplo, en la desintegración $\Sigma^- \to \Sigma^0 + e^- + \nu$ ya considerada, la carga del Σ^- se transfiere al electrón (e). Sin embargo, los procesos en que tal transferencia no tiene lugar, tales como $K^0 \to \bar{e} + e$, $K^0 \to \bar{\mu} + \mu$, que uno esperaría se diera con intensidad $G(\Delta S \neq 0) = G_F \operatorname{sen}\theta \simeq (0.2) \times G_F$, tienen

NOMBRE		SIMBOLO	CARGA	MASA	INTERACCIONES	EXTRAÑEZA	DESINTEGRACION
KAONES	NEUTRO	K^0	0	497,7	s,w	1	$\pi\pi, 3\pi$
	CARGADO	K^+	+1	493,7	s,e.m.,w		$\mu + \nu_\mu$
LAMBDA		Λ	0	1.115,6	s,w	−1	$p + \pi^-$
SIGMA	NEGATIVA	Σ^-	−1	1.197,4	s,e.m.,w	−1	$n + \pi^-$
	NEUTRA	Σ^0	0	1.192,5	s,w		$\Lambda + \gamma$
	POSITIVA	Σ^+	+1	1.189,4	s,e.m.,w		$P + \pi^0$
CASCADAS	NEUTRA	Ξ^0	0	1.314,9	s,w	−2	$\Lambda + \pi^0$
	NEGATIVA	Ξ	−1	1.321,3	s,e.m.,w		$\Lambda + \pi^-$
OMEGA		Ω^-	−1	1.672,3	s,e.m.,w	−3	$\Xi + \pi$

PARTICULAS EXTRAÑAS. Nótense las afinidades y las diferencias con las partículas de la tabla precedente. Cuando estas partículas interaccionan entre sí o con las ordinarias, lo hacen con una fuerza similar a la de Yukawa, pero cuando se desintegran en partículas "ordinarias" en procesos tales como $\Lambda \to p + \pi^-$, $K^- \to \mu + \nu_\mu$, la fuerza de interacción correspondiente varía entre G_F y unas veinticinco veces menor que G_F. Sólo se indican los modos de desintegración más usuales.

NOMBRE (*a*)	SIMBOLO	CARGA	NUMERO CUANTICO INTERNO	MASA (*b*)
UP	u	$^2/_3$	ISOSPIN (*e*)	1,5MeV (*b*)
DOWN	d	$-^1/_3$	ISOSPIN (*e*)	4MeV (*b*)
STRANGE	s	$-^1/_3$	EXTRAÑEZA	200MeV (*b*)
CHARMED	c	$^2/_3$	CHARM (*f*) (ENCANTO)	1500MeV
BOTTOM (*c*)	b	$-^1/_3$	BELLEZA (*f*)	4500MeV
TOP (*d*)	t	$^2/_3$(?)	VERDAD (*f*)	(?)
. ? .		. .	. . . .	. . .

LOS QUARKS son partículas ficticias, introducidas por Gell-Mann y Zweig. *(a)* **Los nombres, en inglés, quieren decir: arriba** *(up)*, **abajo** *(down)*, **extraño** *(strange)*, **encantado** *(charmed)*, **(lo más) bajo** *(bottom)*, **(lo más) alto** *(top)*. *(b)* **La masa de un quark no está bien definida, dado que sólo se pueden encontrar ligados. Esto es particularmente cierto para los quarks más ligeros.** *(c)* **La prueba experimental para este quark es muy reciente, y aun de carácter preliminar.** *(d)* **La prueba para este quark consiste meramente en indicaciones teóricas.** *(e)* **Los quarks** *u* **y** *d* **corresponden cada uno a una de las dos direcciones del espacio de isospín.** *(f)* **Estos nombres, que nos parecen un tanto ridículos, son producto del peculiar sentido del humor de la escuela de Harvard (Estados Unidos).**

lugar con una intensidad que es sólo de $G(\Delta S \neq 0$, neutro$) \simeq 10^{-4}\ G_F$ y recuérdese que la constante de Fermi, G_F, vale 0,0000102 (m_p^{-2}). La explicación de este hecho es uno de los triunfos del modelo del encanto, que consideraremos más adelante.

Si olvidamos por el momento las interacciones débiles y electromagnéticas, tenemos que el esquema de isospín se extiende muy bien a las partículas extrañas *(véase la ilustración de la página 10)*. Pero hay más: puesto que las interacciones fuertes no dependen de la extrañeza, es razonable extender el esquema de isospín para que incluya la extrañeza, aumentando en una la dimensión del espacio interno que así pasa a tener tres dimensiones complejas (dos de isospín y una de extrañeza). El esquema resultante, que en su forma definitiva fue establecido por Gell-Mann y Ne'eman (después de muchos intentos entre los que cabe destacar el modelo de Sakata), tiene como grupo asociado el famoso SU(3). Un hecho salta a la vista en el esquema: a pesar de que las interacciones son independientes del lugar que la partícula ocupa en las representaciones, las masas varían enormemente de una partícula a otra. Por otra parte, y en contra de lo que ocurría con el isospín, cuya representación fundamental (a partir de la cual pueden construirse todas las demás) de dimensión 2 es realizada en la naturaleza (los nucleones), a la representación fundamental de SU(3), de dimensión 3, no corresponden ninguna de las partículas conocidas. Debido a esto, Gell-Mann y Zweig introdujeron unas partículas ficticias (más que como entes reales, como artificio matemático): son los quarks *(véase la ilustración de esta misma página)*, denotados con los símbolos *u*, *d* y *s* (sobre los demás quarks de la tabla hablaré más adelante). La idea es que las partículas conocidas son en realidad agregados de estos entes ficticios, y que sus propiedades pueden entenderse simplemente como expresado su contenido en quarks; por ejemplo, el isospín es simplemente el contenido en quarks *u* y *d*, y la extrañeza el contenido en quarks *s*. Nótese que el modelo de quarks no se aplica a los leptones que, por tanto, siguen considerándose partículas elementales mientras que los hadrones se visualizan como compuestos. La interacción fuerte se explica así como un residuo, a la Van der Waals, de una hipotética fuerza elemental entre quarks *(véanse y confróntense las ilustraciones de las páginas 13 y 9)*.

El esquema de Cabibblo también se simplifica: basta suponer que la partícula que intervienen en las interacciones débiles no es la *d* o la *s*, sino una combinación lineal de las dos:

$$d_\theta = d \cos\theta + s \operatorname{sen}\theta,$$

con lo que la esperada relación $G(\Delta S \neq \neq 0) = G_F \operatorname{sen}\theta \simeq (0.2) \times G_F$ es automática. Las interacciones débiles se supone, pues, que acoplan las partículas en pares que (incluyendo a los leptones) son

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \quad \begin{pmatrix} u \\ d_\theta \end{pmatrix}$$

Como se ve, en todos los acoplamientos hay transferencia de carga eléctrica unidad (del ν_e al e, del u al d_θ, del ν_μ al μ), con lo que, aunque de forma algo forzada, se explica la debilísima intensidad de los procesos $K^0 \to \bar{e} + e$ y $K^0 \to \bar{\mu} + \mu$. Desgraciadamente, el esquema deja fuera a la combinación, ortogonal a d_θ, $s_\theta = -d \operatorname{sen}\theta + s \cos\theta$, lo que no es muy comprensible.

Con cuatro problemas graves se enfrenta, sin embargo, el modelo de quarks "clásico". Estos son: primero, que a pesar de una intensa búsqueda, no hay ninguna prueba convincente de la existencia de quarks aislados. (Hay dos o tres experimentos que han pretendido encontrar quarks; ninguno de ellos ha sido confirmado y lo único que se sabe es que, de existir, los quarks deberían pesar más de cuatro veces lo que el protón, lo que contradice otras indicaciones sobre sus masas.) Segundo, el hecho de que contando cada quark como +1, y cada antiquark como −1, todas las partículas conocidas tienen un número de quarks múltiplo de tres; y además, los quarks se comportan dentro de un protón, por ejemplo, como partículas de tipo Bose-Einstein, cuando, al tener espín 1/2, deberían obedecer el principio de exclusión de Pauli. [El principio de exclusión de Pauli establece que en un mismo sistema no puede haber dos fermiones (partículas cuyo espín es múltiplo impar de 1/2) que sean idénticos en el mismo estado cuántico; a los fermiones se oponen los bosones, que son partículas cuyo espín es entero.] En tercer lugar, las interacciones fuertes entre los quarks (que producen como residuos a la Van der Waals las fuerzas de Yukawa) son independientes del tipo de quark que se considere, *u*, *d* o *s*. Esto, que a primera vista parece muy positivo ya que explica por qué las interacciones entre los hadrones son universales, presenta el problema de que las partículas correspondientes al campo de la interacción elemental entre quarks no puede ser ninguna de las conocidas, ya que todas éstas (debido a sus diferentes masas) violan la necesaria universalidad. Finalmente, aunque las fuerzas que ligan a los quarks tienen que ser muy grandes (de lo contrario sería fácil romper los hadrones y "liberar" a los quarks), cuando se mira un hadrón a cortas distancias (distancias inferiores a unos 10^{-14} centímetros, para lo que se utilizan fotones virtuales) el resultado experimental (funciones de estructura) puede interpretarse muy bien a condición de suponer que los quarks estén prácticamente libres.

Una solución a estas cuestiones fue propuesta por Greenberg en 1924 y, en su forma contemporánea por W. Bardeen, Fritzsch y Gell-Mann. Consistía en suponer que cada quark tiene tres posibles estados que se distinguen por un nuevo número cuántico al que se llamó color. Los colores utilizados dependen de la nacio-

nalidad del autor. Gell-Mann toma siempre los colores de la bandera americana (rojo, blanco y azul); por motivos análogos, muchos físicos españoles utilizan el rojo, amarillo y violeta. Así, por ejemplo, un quark u puede venir en cualquiera de los tres colores, rojo, amarillo o violeta, lo que se denota por u^r, u^a, u^v; y lo mismo el d o el s. Se supone que los quarks interaccionan entre sí por medio de un campo que es el que los liga (al que se ha llamado *gluónico,* del inglés glue=cola de pegar), y que *sólo afecta al color;* la universalidad de las interacciones quark-quark queda pues asegurada. Si se supone que hay ocho de estos gluones, y que las interacciones quark-quark son invariantes bajo cambios de color (lo que nos da un nuevo grupo SU(3), de "color"), puede demostrarse que todos los estados ligados deben tener un número de quarks múltiplo de tres, lo que resuelve la segunda dificultad antes mencionada. La primera y cuarta dificultades no fueron superadas hasta más tarde, cuando Politzer y Gross y Wilczek demostraron que la teoría de color con ocho gluones tiene la propiedad de "libertad asintótica" y, probablemente, "confinamiento". La manera más sencilla de ver esto es escribir el potencial elemental entre dos quarks: éste (compárese con la ley de Coulomb) es del tipo $k_C^2/4\pi$ y multiplicando el resultado por r; en donde r es la distancia entre los quarks.

A cortas distancias r es pequeño ("libertad asintótica") y, por tanto, los quarks aparecen como libres; cuando intentamos separar los quarks, la fuerza crece (r aumenta), lo que explica la dificultad, si no imposibilidad, de aislar quarks ("confinamiento").

Hasta hoy, la inmensa mayoría de las predicciones experimentales del modelo de color (quarks y gluones) están de acuerdo con la experiencia. Salvo un importante problema que mencionaremos en el próximo capítulo, la situación, pues, es relativamente satisfactoria: tenemos una teoría unificada de las interacciones fuertes (QCD: "quantum chromodynamics", nombre por el que se conoce en inglés a la teoría de quarks y gluones con color); una teoría excelente de las interacciones electromagnéticas; y un esquema que, al menos fenomenológicamente, describe las interacciones débiles.

Unificación de interacciones débiles y electromagnéticas. Corrientes neutras. Encanto.

El problema más grave del esquema que acabamos de describir es que el modelo de las interacciones débiles es inconsistente. El Lagrangiano de interacción tipo Fermi (incorporando la teoría Cabibbo y el modelo de quarks) es

$$\mathscr{L}_{\text{int}} = \frac{G_F}{\sqrt{2}} \sum_{\lambda} J_\lambda J_\lambda^+$$

$$J_\lambda = \bar{d}_\theta \gamma_\lambda (1-\gamma_5) u + \bar{e}\, \gamma_\lambda (1-\gamma_5)\nu_e + \bar{\mu}\gamma_\lambda (1-\gamma_5)\nu_\mu$$

donde $d,u,\nu,e,$ ... son los operadores que representan a las partículas correspondientes, y las γ son las matrices de Dirac. Si se intenta calcular con el Lagrangiano de interacción recién formulado más allá del primer orden de teoría de aproximaciones, encontramos resultados infinitos. Esto también ocurre con QED y QCD; mas, para estas últimas, puede demostrarse que la teoría se puede reexpresar de forma que las predicciones para cantidades observables sean siempre finitas. Una teoría con esta propiedad se llama *renormalizable;* otra como la anterior de Fermi, donde siempre se encuentra infinitos, se llama *no-renormalizable;* desde un punto de vista fundamental, es dudoso que una teoría no-renormalizable tenga sentido físico.

Un primer paso para reemplazar el Lagrangiano de interacción tipo Fermi con las incorporaciones citadas por una teoría renormalizable fue dado ya en 1954 por Lee y Yang, quienes postularon la existencia de partículas (las llamadas bosones vectoriales intermedios) W^+, W^-, de gran masa (más de 20 veces la del protón), sustituyendo el tipo Fermi por

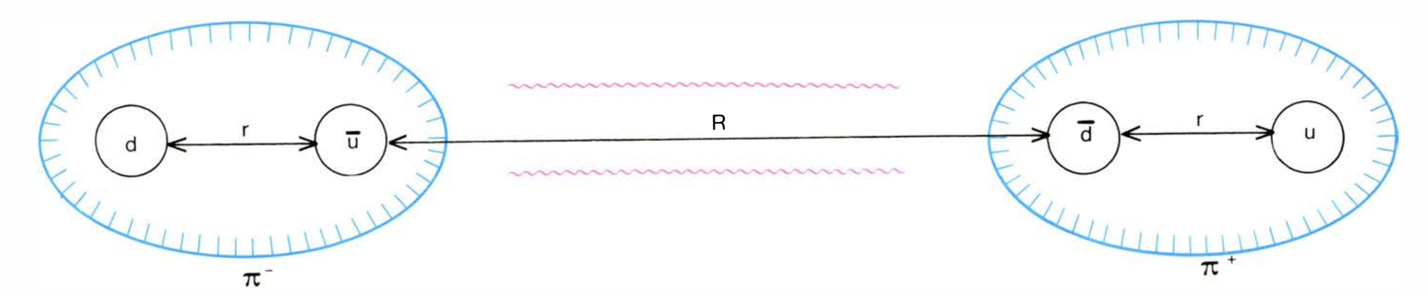
$$\mathscr{L}_{\text{int}} = g_{LY} \Sigma (J_\lambda W_\lambda^+ + J_\lambda^+ W_\lambda^-)$$

Las predicciones son prácticamente las mismas que las del modelo de Fermi con tal que escojamos

$$\frac{g_{LY}^2}{4\pi} = \frac{G_F m_W^2}{4\pi\sqrt{2}} \simeq 10^{-3}$$

Esto es muy interesante, porque el valor de $g_{LY}^2/4\pi$ es del orden de α con lo que aparece una posible relación entre las interacciones débiles y las electromagnéticas. Sin embargo, aunque las divergencias del Lagrangiano de interacción de tipo Lee y Yang son menos violentas que las de tipo Fermi, también aquél sigue siendo no-renormalizable. Para hacerlo normalizable, puede probarse que es necesario que exista un grupo de invariancia interna; las relaciones que esta invariancia fuerza son las que llevan a cancelaciones que hacen desaparecer los infinitos y convierten a la teoría en renormalizable. Esto implica

EL CAMPO (productos de fuerza de interacción), *(en azul),* **se debe a la emisión y absorción virtuales de la partícula P asociada, por dualidad cuántica, a dicho campo. El alcance de la fuerza es** *d.*



FUERZAS entre dos piones como residuo de las interacciones elementales de las partículas (quarks *u, d*) que los componen. La semejanza con las fuerzas que actúan entre los átomos de la molécula de hidrógeno antes expuesta *(véase la ilustración de la página 9)* **resulta evidente.**

la existencia de un nuevo bosón vectorial neutro, W^o y añadir al Lagrangiano de interacción de tipo Lee y Yang las interacciones siguientes *(corrientes neutras):*

$$g_{LY}\sum_{\lambda}(j^{l}_{\lambda} W^{o}_{\lambda} + j^{h}_{\lambda} W^{o}_{\lambda}), \text{ donde:}$$

$$j^{l}_{\lambda} = \frac{1}{\sqrt{2}}\{\bar{\nu}_e \gamma_{\lambda} (1-\gamma_5) \nu_e - \bar{e} \gamma_{\lambda} (1-\gamma_5) e + $$
$$+ \bar{\nu}_{\mu} \gamma_{\lambda} (1-\gamma_5)\nu_{\mu} - \bar{\mu} \gamma_{\lambda} (1-\gamma_5)\mu\}, \text{ y}$$

$$j^{h}_{\lambda} = \frac{1}{\sqrt{2}}\{\bar{d}_{\theta} \gamma_{\lambda} (1-\gamma_5) d_{\theta} - \bar{u} \gamma_{\lambda} (1-\gamma_5) u\}$$

El trozo que contiene a la expresión que define a j^{l}_{λ} no había sido observado, pero ciertamente esto podría deberse a dificultades experimentales: el trozo que contiene a la expresión que define j^{h}_{λ} induce, sin embargo, corrientes neutras con cambio de extrañeza y desintegraciones, como las $K^o \to \bar{e} + e$, $K^o \to \bar{\mu} + \mu$, con intensidad similar a las de un proceso débil típico; y ello está en violento desacuerdo con la experiencia. Sin embargo, este problema sólo aparece para los hadrones; por tanto, se puede pensar en construir un modelo válido sólo para los leptones; es decir, introducir sólo las interacciones de la expresión que define a j^{l}_{λ}, pero no las de la expresión que define a j^{h}_{λ}, aunque esto sea inconsistente.

Avanzando en esta línea, el parecido entre $g^2_{LY}/4\pi$ y $e^2/4\pi = \alpha$ hace natural, como se ha dicho, el intentar unificar las interacciones débiles y electromagnéticas. En las primeras unificaciones (Schwinger, Glashow [1961]) se introducía un nuevo campo neutro, *B*, de forma que éste y W^o se mezclan para dar dos partículas físicas: la llamada *Z*, muy masiva, y el fotón, γ, de masa cero:

$$Z = W^o \cos\theta_W + B \operatorname{sen}\theta_W$$
$$\gamma = W^o \operatorname{sen}\theta_W + B \cos\theta_W$$

Al ángulo θ_W se le llama *ángulo de mezcla débil.*

El modelo tal cual tiene el indudable atractivo de unificar las interacciones electromagnéticas y débiles; pero debido a la diferencia de masas entre el *Z*, los W^+ y W^- y el fotón, la simetría interna, conocida por "isospín débil", se rompe, y el modelo deja de ser renormalizable: por lo que estas teorías cayeron en el olvido hasta ser resucitadas por Weinberg (1967) y Salam (1968) al descubrirse cómo aplicar al problema el mecanismo de ruptura espontánea de simetría, que paso a discutir brevemente.

Tomemos un ejemplo clásico. Piénsese en un vaso y una bolita *(véase la ilustración de la página 14).* El vaso tiene, claramente, simetría de revolución alrededor del eje *Z*. Sin embargo, hay dos posiciones de equilibrio de la bolita: la posición I, que sí es simétrica, y la II, que no lo es. La simetría inicial de revolución ha desaparecido; únicamente se manifiesta en el hecho de que en adelante no habrá una única posición de equilibrio, sino una serie de ellas (las II', II"... en la ilustración) entre todas las cuales recomponen la simetría inicial.

Estos mecanismos de ruptura espontánea de la simetría son conocidos desde hace mucho tiempo; aplicados a la física de partículas reciben el nombre de *mecanismo de Higgs,* quien fue uno de los físicos que lo introdujeron. Un aspecto de interés incluso filosófico de este mecanismo en el caso de la mecánica cuántica relativista es que, aquí, el estado de equilibrio estable es el de menor energía, y corresponde al vacío. Por tanto, el hecho de que las soluciones con ruptura espontánea de simetría no sean únicas implica que el vacío tampoco es único; precisamente, la simetría entre los distintos vacíos posibles se rompe al considerar a uno de ellos como el vacío físico.

De todas maneras, tampoco el modelo de Weinberg-Salam se tomó muy en serio, hasta que en 1971 't Hooft demostró que una teoría con simetría interna de tipo *gauge,* rota espontáneamente, es de hecho, renormalizable. Antes de seguir, sin embargo, es importante dedicar unas palabras al calificativo *gauge,* que acabamos de utilizar. Se dice que una simetría es de tipo *gauge local,* o simplemente *gauge,* si se tiene invariancia incluso cuando las rotaciones que se hacen en el espacio interno varían según el punto del espacio ordinario en que se realizan. Para que se dé esta invariancia, puede demostrarse que es necesario que todas las interacciones que corresponden a la misma invariancia tengan la misma intensidad. El poder unificador de este hecho es enorme: quiere decir que el que una interacción tenga una intensidad constante, independiente de las partículas específicas que se consideren, no es ninguna hipótesis ad hoc, sino una consecuencia de la teoría (invariancia *gauge*). Debido a esta capacidad unificadora, y al hecho de garantizar renormalizabilidad, pocos físicos dudan que la teoría fundamental sea de tipo *gauge.*

Cerrada esta disquisición, digamos que a pesar de los atractivos extra de que el descubrimiento de 't Hooft dotó al modelo unificado de interacciones débiles y electromagnéticas descrito, queda por resolver el problema de los nuevos términos j^{l}_{λ} y j^{h}_{λ} que hemos tenido que introducir. Fijándonos en los términos que sólo involucran leptones j^{h}_{λ}, estas interacciones implican la existencia de corrientes neutras, en particular, de procesos como las colisiones: $\nu_{\mu} + e \to \nu_{\mu} + e$, con intensidad análoga a la de los procesos con corrientes cargadas, es decir, con intercambio de carga eléctrica, del tipo $\nu_{\mu} + n \to \mu + p$.

Después de una búsqueda intensa, estos procesos con corrientes neutras fueron encontrados primero en el CERN y después en otros laboratorios. La intensidad era la predicha teóricamente. El ángulo de mezcla interacción débil-interacción electromagnética (θ_W), introducido antes, fue medido experimentalmente. Su valor es tal que $\operatorname{sen}^2\theta_W = 3{,}5 \pm 0{,}6$.

Este éxito predictivo de la teoría unificada de interacciones débiles y electromagnéticas no hizo sino resaltar aún

más el ya citado problema de las corrientes neutras con cambios de extrañeza. Es fácil ver que no hay manera de suprimirlas dentro del esquema con tres quarks, si bien no resulta difícil hacerlo si se postula un nuevo quark. El esquema más sencillo en el que esto es posible, debido a Glashow, Iliopoulos y Maiani, introduce un nuevo quark, c (quark encantado), que lleva un nuevo número cuántico (encanto, "charm" en inglés). Las interacciones fuertes de este quark son, como es lógico, las mismas que las de los demás; al igual que la extrañeza, se supone que el encanto sólo puede crearse o destruirse por medio de las interacciones débiles en las que intervienen con el quark que, hasta ahora, habíamos dejado de lado, el s_θ; por consiguiente, las interacciones débiles que acoplan las partículas en pares serán

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} u \\ d_\theta \end{pmatrix}, \begin{pmatrix} c \\ s_\theta \end{pmatrix}.$$

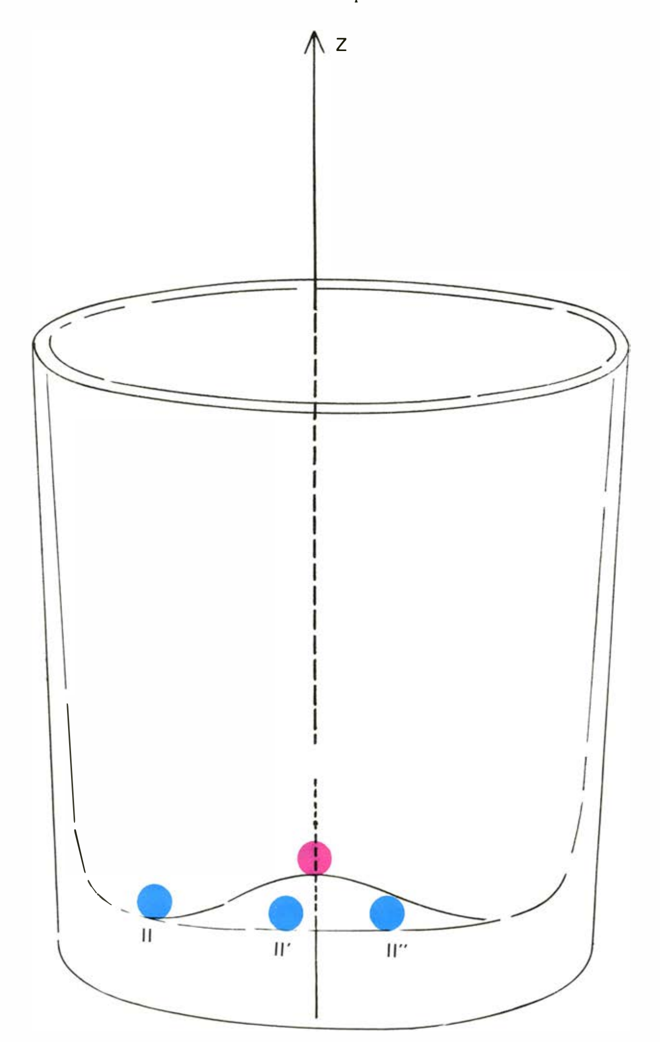
El carácter gauge de la simetría implica que hay que añadir un nuevo término a las interacciones débiles. Un término $\bar{s}_\theta \gamma_\lambda (1-\gamma_5) c$ a J_λ en el Lagrangiano de interacción antes expuesto, y un término

$$\frac{1}{\sqrt{2}} \{\bar{s}_\theta \gamma_\lambda (1-\gamma_5) s_\theta - \bar{c} \gamma_\lambda (1-\gamma_5) c\},$$

a j^{h}_λ. *No es difícil ver que este último añadido cancela precisamente a los términos indeseables* (corrientes neutras con cambio de extrañeza) de la interacción j^{h}_λ.

Gráficamente podemos imaginar que un proceso como $K^0 \to \mu + \bar{\mu}$ está solicitado por dos fuerzas; una, proviniente de j^{h}_λ, de intensidad $d\, \gamma_\lambda (1-\gamma_5) \bar{s} \cos\theta \,\mathrm{sen}\, \theta$, y otra, debido a la interacción definida en el punto y aparte anterior, de valor $-d\gamma_\lambda (1-\gamma_5) s \cos\theta \,\mathrm{sen}\, \theta$, que la anula.

De hecho la cancelación no es exacta. Puede demostrarse (Glashow, Iliopoulos y Maiani; B. Lee y M. Gaillard) que quedan efectos residuales de orden $(m_c^2 G_F)^2$. Dada la probabilidad experimental del proceso $K \to \mu + \bar{\mu}$, resulta la predicción de unos 1500 MeV para la masa del quark encantado, lo que a su vez implica una masa entre 2000 y 4000 MeV para las partículas físicas que contengan quarks encantados.



POSICIONES DE EQUILIBRIO de una bolita en el fondo de un vaso. Posición simétrica I *(en rojo)* **inestable. Posiciones asimétricas II, II' II''** *(en azul)* **estables. Entre todas éstas recomponen la simetría de rotación alrededor del eje Z. La simetría se rompe al particularizar una de ellas.**

El esquema fue confirmado brillantemente por la experiencia. En 1975, Richter y Ting encontraron las partículas ψ, que contienen encanto oculto (es decir, son estados ligados de encanto y antiencanto), y, posteriormente, se han ido describiendo partículas con encanto aparente (estados ligados de quarks con encanto y sin encanto). Como dijimos en el comienzo de este trabajo, son estas dos verificaciones experimentales (corrientes neutras y descubrimiento del encanto) las que han dado apoyo a los modelos unificados de interacciones débiles y electromagnéticas.

Gran unificación, superunificación. Gravitación. Situación actual

Recapitulemos la situación tal como la acabamos de describir. Las interacciones fuertes vienen descritas por una simetría gauge asociada con el grupo SU(3)' que actúa sobre el color de los quarks. Las interacciones débiles y electromagnéticas se unifican en una interacción, también de tipo gauge, basado en el grupo SU(2)X U(1). Las interacciones fuertes vienen mediadas por un campo cuyas partículas asociadas (los gluones) no son observables aisladamente por estar, igual que los quarks, confinados dentro de los hadrones "físicos". Las interacciones débiles y electromagnéticas están mediadas por un campo con cuatro partículas asociadas: W^+, W^-, Z, y el fotón, γ. Las tres primeras son muy pesadas (unos setenta a cien veces la masa del protón) y no se ha observado experimentalmente aún porque no se dispone de energía suficiente para crearlas; el fotón no tiene masa. Esta ruptura de simetría es del tipo espontáneo (mecanismo de Higgs). Lo único que no está fijado en el esquema es el número de quarks y leptones.

Cuando todo el mundo estaba contento con los cuatro leptones "clásicos" y los cuatro quarks u, d, s y c, se ha descubierto un nuevo leptón τ de masa unos 2000 MeV (¡el nombre de "leptón" empieza a no ser apropiado!) y, posiblemente, su neutrino asociado, ν_τ; y hay fuertes indicios de la existencia de un nuevo quark, b *(véase la ilustración de la página 11)*, lo que probablemente implica que existe al menos otro t, aún más pesado. La existencia de tales partículas no afecta al esquema que aparece, casi por primera vez desde el siglo XIX, como un modelo global, capaz de dar una descripción unificada de las interacciones básicas.

Sin embargo, parece difícil evitar el pensar que las interacciones fuertes y las electromagnético-débiles no sean, de

nuevo, sino dos aspectos distintos de una única interacción superfundamental. En efecto, ambas interacciones tienen mucho en común: ambas son mediadas por bosones vectoriales y ambas son de tipo gauge. Su distinta intensidad no es un problema: puede demostrarse que, a grandes energías, ambas intensidades serán comparables, con lo que su diferencia resultaría, de nuevo, un caso de simetría espontáneamente rota.

Tal unificación ("gran unificación", en la jerga del oficio) ha sido propuesta por Pati y Salam, y posteriormente por muchos otros físicos. De hecho, estos esquemas responden a una vieja aspiración que en cierto sentido, y como hemos visto, se remonta a la filosofía presocrática; lo que es nuevo es que se llega a implicaciones específicas, que se pueden refrendar con la experiencia. Por ejemplo, los modelos predicen el ángulo de mezcla débil (en el modelo de Georgi y Glashow, $\text{sen}^2\theta_W=3/8$), predicción que está en acuerdo con el valor experimental, valor tal que $\text{sen}^2\theta_W=3{,}5\pm0{,}6$; y que no hay interferencia débil-electromagnética (C. Jarlskog, Ynduráin [1976]) en aniquilaciones inclusivas electrón-positrón, lo que es compatible con resultados experimentales actuales y podrá ser verificado con precisión en experimentos que se harán pronto. A pesar de estas indicaciones de que la idea de una unificación global es correcta, lo cierto es que no hay consenso sobre el modelo concreto en que plasmarlo; para distinguir entre las distintas alternativas se necesita mayor información experimental.

Hemos venido pasando revista a una serie de modelos, en los que, progresivamente, se han ido unificando más y más interacciones. En todos los casos, los esquemas están bien establecidos en el sentido de que la mayor parte de los expertos en el tema los consideran como mecanismos apropiados para describir los procesos en cuestión, y hay pruebas experimentales en su corroboración. No podemos, sin embargo, acabar este artículo sin mencionar al menos esquemas de carácter más especulativo, que, no obstante, pretenden ir más allá aún: dedicaremos pues este apartado final a discutir problemas abiertos y a los esfuerzos por resolverlos.

En primer lugar, incluso lo que llamamos "gran unificación" contiene dos tipos de partículas: quarks y leptones, de espín 1/2, y gluones, W, Z, γ, de espín 1. Antes de que se valorase la importancia de las simetrías gauge, se intentó construir modelos en los que las partículas de espín 1 eran estados ligados de las de espín 1/2 (Heisenberg). Sin embargo, el carácter gauge de las simetrías implica que las partículas de espín 1 deben ser elementales. Nos quedan, pues, dos posibilidades: construir las partículas de espín 1/2 a partir de las de espín 1, lo que, sorprendentemente, es posible ('t Hooft, 1976); o intentar unificar partículas de distintas especies en una supersimetría (Wess y Zumino, 1973). La primera posibilidad no se sabe bien cómo implementarla; la segunda, por lo menos en su forma actual, contiene predicciones en desacuerdo con la experiencia.

Un segundo problema importante es el intentar predecir, y no simplemente introducir como parámetros fenomenológicos, las masas de las partículas, y los grupos de invariancia. La única masa que puede formarse con unidades naturales es la de Plank:

$$m_{\text{Pk}}=\sqrt{\frac{\hbar c}{G_{\text{N}}}}$$

que contiene la constante de Newton, G_{N}; lo que sugiere que no se puedan calcular masas sin tener en cuenta la interacción gravitatoria (que hasta ahora hemos dejado al margen). Hay otro motivo por el que se insinúa una conclusión similar. En las grandes unificaciones, la ruptura espontánea entre interacciones fuertes y débiles-electromagnéticas se da a una energía E del orden de (Georgi, Quinn y Weinberg, 1974).

$$E \sim (1\,\text{GeV}) \log\left(\frac{g_G}{e}\right)$$

donde e es la carga del electrón y g_G la constante de acoplamiento gluónica (a 1 GeV). Los valores que resultan para E dependen del modelo específico pero son comparables a m_{P}.

Finalmente, tenemos el problema de qué hacer con la interacción gravitatoria. La interacción gravitatoria es, cuando se cuantifica, una teoría gauge; sin embargo, no se sabe aún si es o no renormalizable. Con respecto a esto, sin embargo, hay fuertes indicaciones de que, para ser renormalizable, es necesario que la materia se ordene según esquemas de simetría similares a los que ya hemos descrito. De hecho, el problema con estas teorías supersimétricas de la gravitación (en las que se da un significado geométrico no sólo a la propia gravitación, como en la teoría original de Einstein, sino también a los fermiones– "quarks" y "leptones") es que unifican *demasiado:* no hay, de momento al menos, ningún modelo de los que funcionan fenomenológicamente que cumplan todas las limitaciones requeridas. De todas maneras, los esquemas de "supergravedad" son tremendamente atractivos puesto que en ellos la unificación de distintas interacciones sería total.

La cuestión del dióxido de carbono

La actividad humana está incrementando notablemente el contenido en dióxido de carbono de la atmósfera terrestre: ¿Podrán los bosques y los océanos retener una cantidad suficiente de carbono que evite un cambio importante en el clima?

George M. Woodwell

Desde 1850, y en un período de 125 años, las actividades humanas han incrementado la cantidad de dióxido de carbono existente en la atmósfera terrestre desde 290 partes por millón o menos a poco más de 330 partes por millón. Posiblemente, la cuarta parte de este aumento ha tenido lugar en el transcurso de la última década. Si la tendencia actual continúa, la cantidad de dióxido de carbono atmosférico podría, en el año 2020, doblar los valores actuales. Hasta fecha muy reciente, este incremento era comúnmente atribuido a la combustión de combustibles fósiles. Hoy disponemos de pruebas que confirman su atribución, en igual grado, a otra causa: la destrucción de los bosques en todas las regiones del mundo.

Aunque el dióxido de carbono es un gas que se halla únicamente a nivel de trazas en la atmósfera terrestre (la concentración a que se encuentra es tan sólo de alrededor de un 0,03 por ciento del volumen), desempeña posiblemente un papel crítico en el control del clima terrestre por absorber la energía radiante en la longitud de onda del infrarrojo. El calor atrapado de esta forma tiene una gran capacidad potencial de alterar sustancialmente el clima mundial. Aparte de sus posibles efectos sobre el clima, el dióxido de carbono desempeña en la atmósfera un papel crucial como fuente del carbono que es fijado en la fotosíntesis por las plantas verdes y que provee, por tanto, la base para toda la vida animal y vegetal.

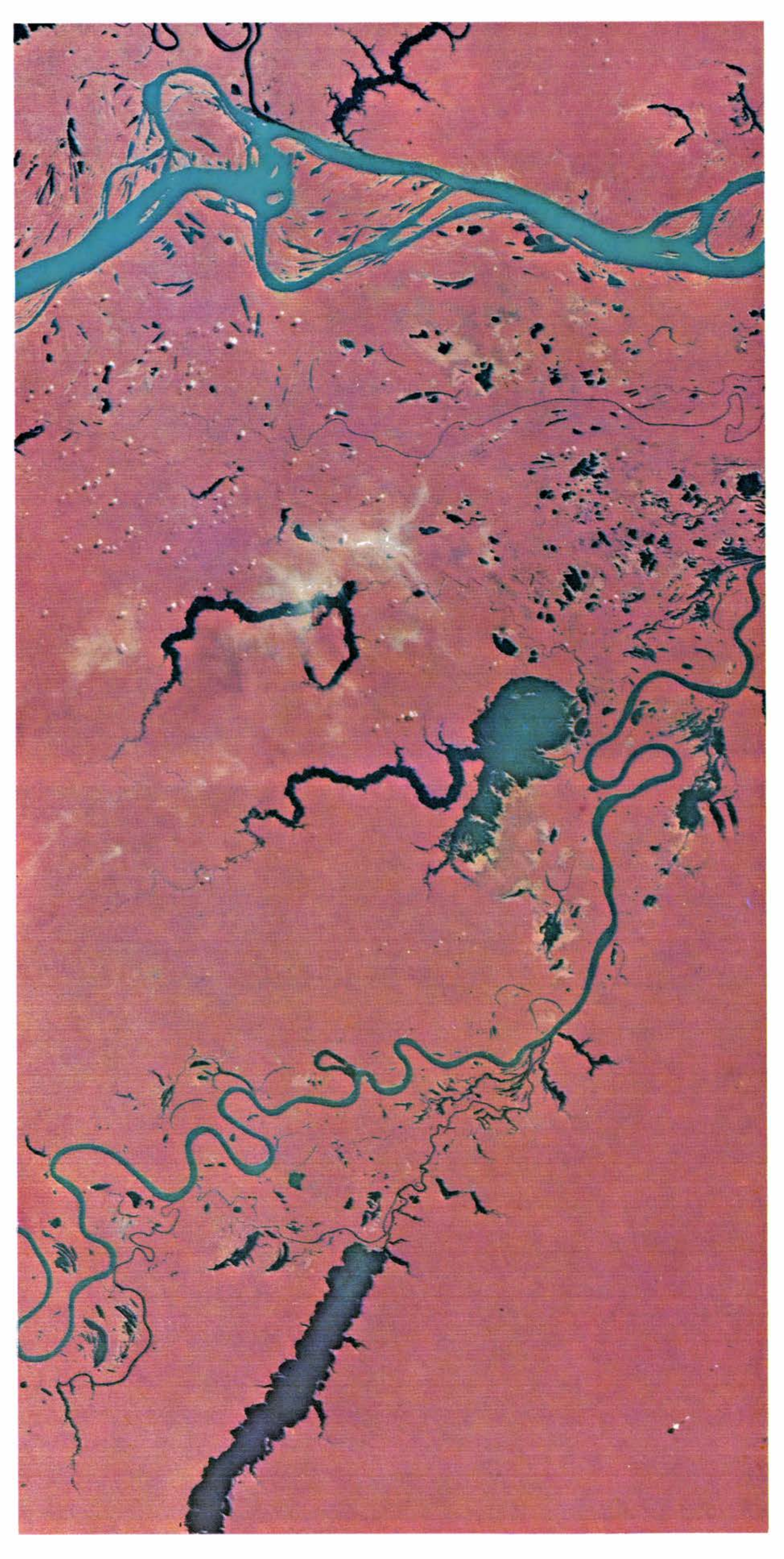
Consecuentemente, la humanidad se enfrenta ante un dilema histórico. Las actividades humanas que están incrementando el contenido en dióxido de carbono de la atmósfera presagian un calentamiento general del clima en las próximas décadas. Aunque no se puede saber con certeza, ni el grado en que cambiará el clima, ni los mecanismos que se verán implicados, los resultados del continuo aumento de la cantidad de dióxido de carbono en la atmósfera serán, casi con seguridad, desestabilizadores. Un incremento de las temperaturas medias mundiales aumentaría probablemente la extensión de las zonas áridas y afectaría de forma significativa a la producción agrícola.

El otro miembro del dilema alude a lo siguiente: las acciones correctoras que cabe esperar que se adopten tendrían a buen seguro efectos igualmente desestabilizadores. La acción correctora más elemental sería la de reducir de forma notable el consumo de combustibles fósiles. De igual importancia serían las medidas encaminadas a retardar el ritmo que se está manteniendo en la reducción de los bosques del mundo por las actividades madereras, la extensión de las tierras agrícolas y de pasto, la intoxicación y otras consecuencias del desarrollo industrial. Un esfuerzo importante encaminado a cambiar el equilibrio del uso del suelo entre agricultura y silvicultura, sumado a una acción destinada a restringir la combustión de combustibles fósiles, establecería patrones de desarrollo económico y social equivalentes a los drásticos cambios que en las condiciones de la vida humana podría producir un calentamiento del clima.

Aunque el problema del dióxido de carbono ha coexistido con nosotros durante más de un siglo, los datos inequívocos de los cambios del contenido en dicho gas en la atmósfera tan sólo se han registrado desde 1958. En dicho año, Charles D. Keeling, de la Scripps Institution of Oceanography, emplazó una estación de análisis continuo de dióxido de carbono en el volcán Mauna Loa, de la isla de Hawaii. Se escogió el Mauna Loa porque ofrecia la oportunidad de estudiar el contenido en dióxido de carbono de la mezcla de aire de la troposfera, o baja atmósfera, en latitudes medias. Los registros que se han venido obteniendo en Mauna Loa y en otras estaciones muestran dos claros patrones de conducta. El primero de ellos ha sido que, desde 1958 y durante todo el período estudiado, ha existido, de forma regular, una tendencia al alza en el contenido en dióxido de carbono atmosférico. La cantidad de este aumento en el Mauna Loa ha sido de alrededor de 0,8 partes por millón por año, aunque lógicamente han existido variaciones en la tasa de incremento. El segundo patrón de conducta apreciado es que existe una oscilación sistemática del contenido en dióxido de carbono de la atmósfera, oscilación que viene correlacionada con las estaciones climáticas. El contenido en dióxido de carbono alcanza un valor máximo a finales de invierno, usualmente abril en el hemisferio norte, y desciende a un mínimo a finales de verano, en los últimos días de septiembre o en octubre. Los datos recogidos en Mauna Loa son los registros continuos de las concentraciones de dióxido de carbono, de mayor duración y precisión que nunca se hayan hecho en el mundo. Correspondientes a varios períodos se han medido y registrado las concentraciones de dióxido de carbono en el Polo Sur, Australia, Punta Barrow en Alaska, Long Island en Nueva York y en otros lugares. Con la ayuda de aeroplanos, los investigadores han sacado amplias muestras de la atmósfera en Estados Unidos, Suecia, Australia y otros lugares. Todos los datos ponen de manifiesto una variación invierno-verano con un mínimo a finales de verano y un máximo a finales de



IMAGENES DE VERANO Y OTOÑO, tomadas por el satélite Landsat, de una región a lo largo del borde este de las Rocosas cerca de Boulder, Colorado; muestran el pulso anual de la fijación del dióxido de carbono que es característica de las zonas templadas. En este tipo de imágenes, el color verde de la vegetación aparece rojo. En la fotografía superior, realizada en agosto, el intenso color rojo del área situada a la derecha de las montañas de la izquierda refleja un valor máximo de actividad fotosintética, en el cual las plantas recaban de la atmósfera el dióxido de carbono para su crecimiento. En la fotografía inferior, realizada en octubre, el color rojo ha desaparecido sustancialmente, lo que pone de manifiesto una disminución de la actividad fotosintética. En esta región la actividad se debe a los bosques y a las tierras de cultivo.



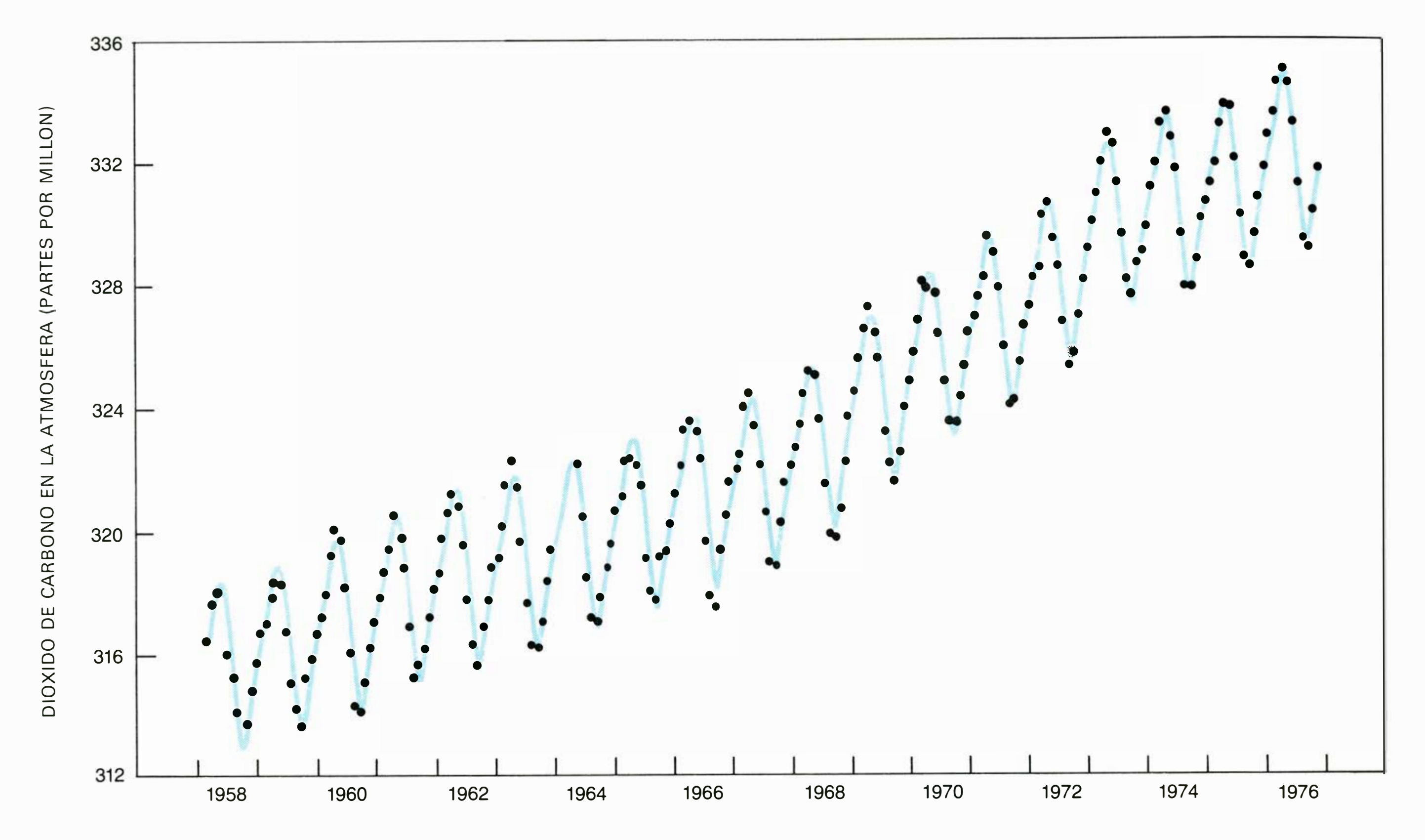
EL AREA DE BOSQUES DE LLUVIA de la Cuenca Amazónica del noroeste de Brasil aparece en esta foto, tomada desde satélite, casi completamente roja, reflejando la intensa actividad fotosintética durante todo el año, que caracteriza a los bosques de lluvia tropicales. Las comunidades naturales de bosques fijan más carbono por unidad de superficie que la mayor parte de las tierras de cultivo agrícola. La extensión de estos bosques está siendo reducida por la mano del hombre.

invierno. La oscilación se correlaciona con las estaciones climáticas de cada hemisferio. Los datos obtenidos muestran también un mayor o menor incremento continuado del contenido en dióxido de carbono de la atmósfera, aumento que varía con el tiempo y con el lugar entre 0,5 y 1,5 partes por millón y por año, aproximadamente.

El cambio estacional en el dióxido de carbono atmosférico refleja uno de los más importantes factores que afectan a la atmósfera: el metabolismo de los biota o de la totalidad de la materia viva. El cambio estacional en las concentraciones de dióxido de carbono en la atmósfera está correlacionado con el "pulso" de la fotosíntesis que tiene lugar durante el verano en las latitudes medias de ambos hemisferios. Recientemente se ha reconocido con la máxima probabilidad que la causa principal de este cambio estacional es el pulso de la fotosíntesis en los bosques de las latitudes medias. El énfasis se centra en los bosques, pues ocupan una superficie muy extensa, realizan más fotosíntesis que cualquier otro tipo de vegetación y tienen la capacidad de almacenar carbono en cantidades suficientemente grandes como para afectar al contenido en dióxido de carbono de la atmósfera.

La variación en la amplitud de la diferencia entre la concentración invernal de dióxido de carbono y la concentración de finales de verano concuerda con esta hipótesis. La diferencia oscila entre unas 5 partes por millón,en Mauna Loa, a más de 15 partes por millón en la zona central de Long Island. La diferencia desciende en los Trópicos, donde el pulso estacional del metabolismo es menos pronunciado o no existe; la diferencia es también menor cuanto mayor es la altura, en todas las latitudes. La amplitud es sustancialmente menor en el hemisferio Sur, a causa, según parece, de la menor masa continental de tierra, lo que limita la superficie de bosques. La conclusión clara que se obtiene es la de que los bosques de la tierra tienen una influencia decisiva en el contenido, a corto plazo, de dióxido de carbono en la atmósfera.

La razón del incremento a largo plazo del contenido en dióxido de carbono del aire (un incremento del 10 al 15 por ciento desde 1850) se ha venido asignando al vertido de dióxido de carbono a la atmósfera procedente de la combustión de combustibles fósiles, a un ritmo acelerado. Pero los análisis recientes han puesto en entredicho esta hipótesis. Mis colegas y yo, en el Ecosystems Center del Marine Biological Laboratory en Woods Hole, Mass., en colaboración



TENDENCIA del contenido en dióxido de carbono de la atmósfera, medida desde 1958 en el observatorio de Mauna Loa, de la isla de Hawaii, por Charles D. Keeling, de la Scripps Institution of Oceanography. Los puntos indican la concentración media mensual de dióxido de carbono. Las oscilaciones estacionales vienen causadas por la captación de dióxido de carbono por el proceso fotosintético que tiene lugar durante la estación de crecimiento en el hemisferio Norte y por la subsiguiente liberación del mismo durante los meses de otoño e invierno. Las mediciones realizadas en Mauna Loa, así como las efectuadas en otros emplazamientos, muestran que el contenido medio de dióxido de carbono en la atmósfera ha aumentado en más de un 5 por ciento desde 1958. La tasa de aumento ha ido variando de año en año por causas todavía desconocidas. La tasa actual de crecimiento es de una parte por millón al año, lo que resulta equivalente a $2,3 \times 10^{15}$ gramos de carbono.

con R. H. Whittaker y Gene E. Likens, de la Cornell University, W. A. Reiners, de Dartmouth College, y C. C. Delwiche, de la Universidad de California en Davis, hemos demostrado que probablemente existe un considerable desprendimiento adicional de dióxido de carbono procedente de los biota. Otros investigadores, en particular Bert Bolin, de la Universidad de Estocolmo, y J. R. Adams, y sus colegas en la Rice University, han dado cuenta de hallazgos similares. El desprendimiento de dióxido de carbono por los biota viene producido principalmente por la destrucción de los bosques y la oxidación del humus. La suposición de que el incremento del contenido en dióxido de carbono de la atmósfera se debía a la combustión de combustibles fósiles, sin contemplar los posibles cambios de los biota, ha conducido a lo que hoy en día parece ser un error de bulto en el cálculo del balance mundial del carbono.

La dificultad surgió de la presunción de que los biota separaban y retenían el dióxido de carbono, cuando en realidad han sido probablemente una fuente de emisión de dióxido de carbono a la atmósfera. Este probable error ocasiona que por el momento no sea posible resolver los principales problemas que atañen al balance mundial del carbono. Si los biota no son un reservorio del dióxido de carbono atmosférico, y si la absorción de dicho gas por los océanos no es mayor de lo que se creía, la cantidad de dióxido de carbono en la atmósfera debería incrementarse de una forma aún más veloz de lo que las observaciones demuestran. Obviamente, estas apreciaciones resultan erróneas. Pero, ¿dónde está el error?

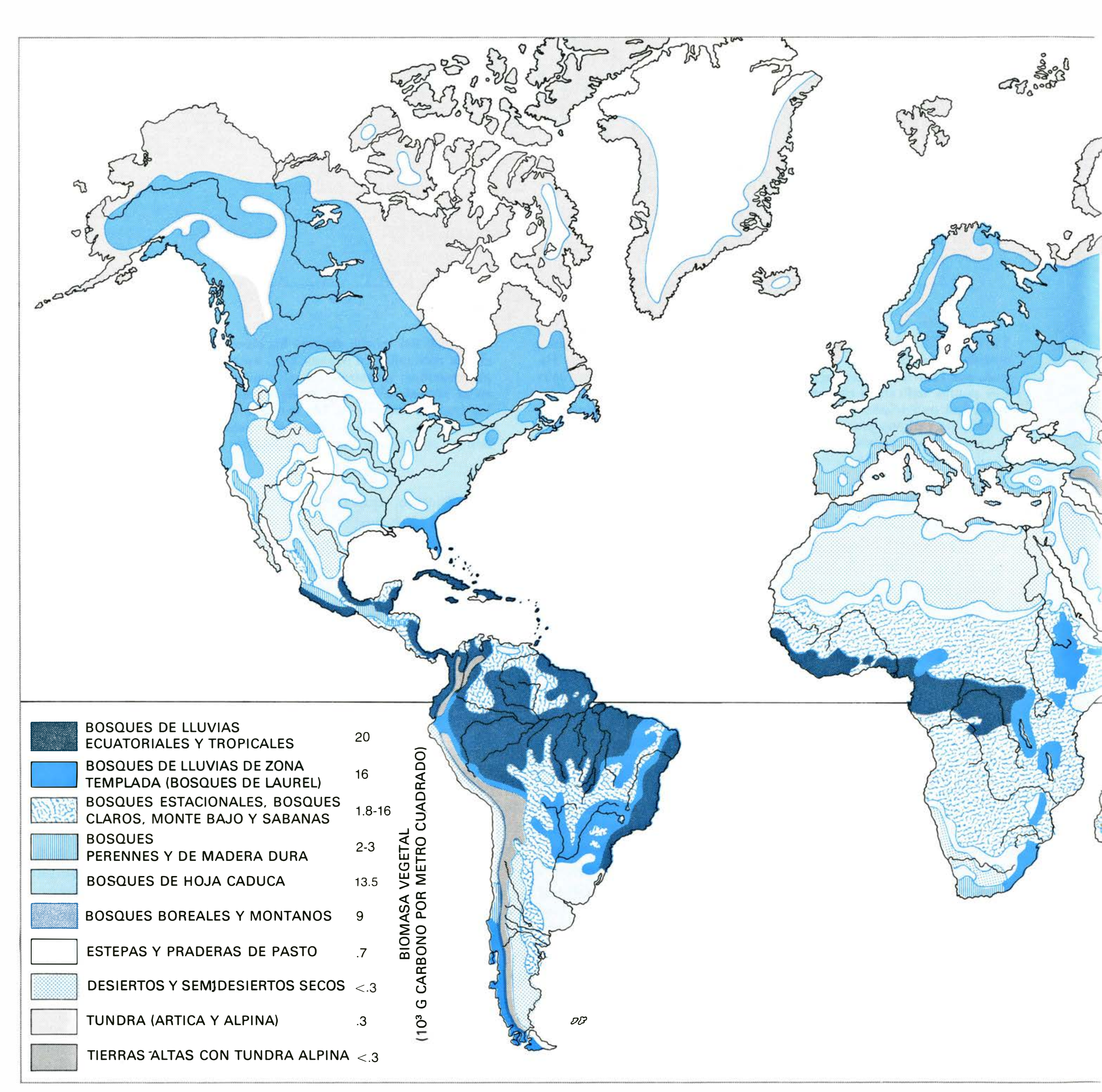
El tema puede exponerse con mayor claridad mediante la comparación de las magnitudes de los diversos sistemas que están de forma más o menos continua intercambiando entre sí carbono. En la actualidad, la atmósfera contiene alrededor de 700×10^{15} gramos de carbono en forma de dióxido de carbono, el cual se está intercambiando incesantemente con los biota y con las aguas superficiales del océano. La cantidad de carbono encerrada en el conjunto de los biota de todo el mundo es aproximadamente de 800×10^{15} gramos, o sea, algo más de lo contenido en la atmósfera. Una cantidad aún mayor de carbono, de un valor estimado entre 1000×10^{15} y 3000×10^{15} gramos, se halla en la materia orgánica del suelo, principalmente humus y turba. La tala de los bosques, la extensión de la agricultura sobre suelos con grandes cantidades de materia orgánica y la destrucción de las tierras húmedas aceleran la desaparición del humus, que se transforma en dióxido de carbono, agua y calor. El dióxido de carbono así emitido entra a formar parte de la atmósfera.

Aunque estos tres depósitos de carbono, en continua interacción, son aproximadamente de la misma capacidad, la cantidad total de carbono contenida en los océanos es mucho mayor. Casi todo el carbono está presente en forma de dióxido de carbono disuelto, el cual es parte del sistema carbónico-carbonatos. La cantidad total de esta reserva se aproxima a los 40.000×10^{15} gramos, incluidas las aguas de los grandes océanos. En una escala de tiempo de miles de años, el contenido en dióxido de carbono de la atmósfera podría ser bien determinado por el equilibrio establecido con el carbono inorgánico de las profundidades oceanicas. Sin embargo, la tasa de intercambio entre la atmósfera y los océanos como un todo es baja. Los intercambios más rápidos tienen lugar entre la atmósfera y la capa su-

perficial de mezcla, aproximadamente los 100 metros superiores, por encima de las frías aguas abisales. La capa superficial contiene unos 600 × 10^{15} gramos de carbono inorgánico. El segundo mayor reservorio de carbono en los océanos es la materia orgánica disuelta (el "humus" del océano) que parece estar por doquier en la proporción de una parte por millón y que puede totalizar hasta 3000 × 10^{15} gramos en los océanos, considerados globalmente.

Las aguas profundas del océano exceden ampliamente en volumen a las aguas de la capa mezclada y contienen de largo la mayor cantidad de carbono que está en intercambio con la atmósfera: de 35.000 × 10^{15} a 38.000 × 10^{15} gramos (se excluye de esta cifra el carbono de los sedimentos carbonáceos, que contienen una cantidad mucho mayor). La capacidad que detentan las regiones abisales de absorber carbono es virtualmente ilimitada. El problema estriba en que, al parecer, el carbono se desplaza muy lentamente desde la atmósfera a las profundidades oceánicas a través de la capa de mezcla.

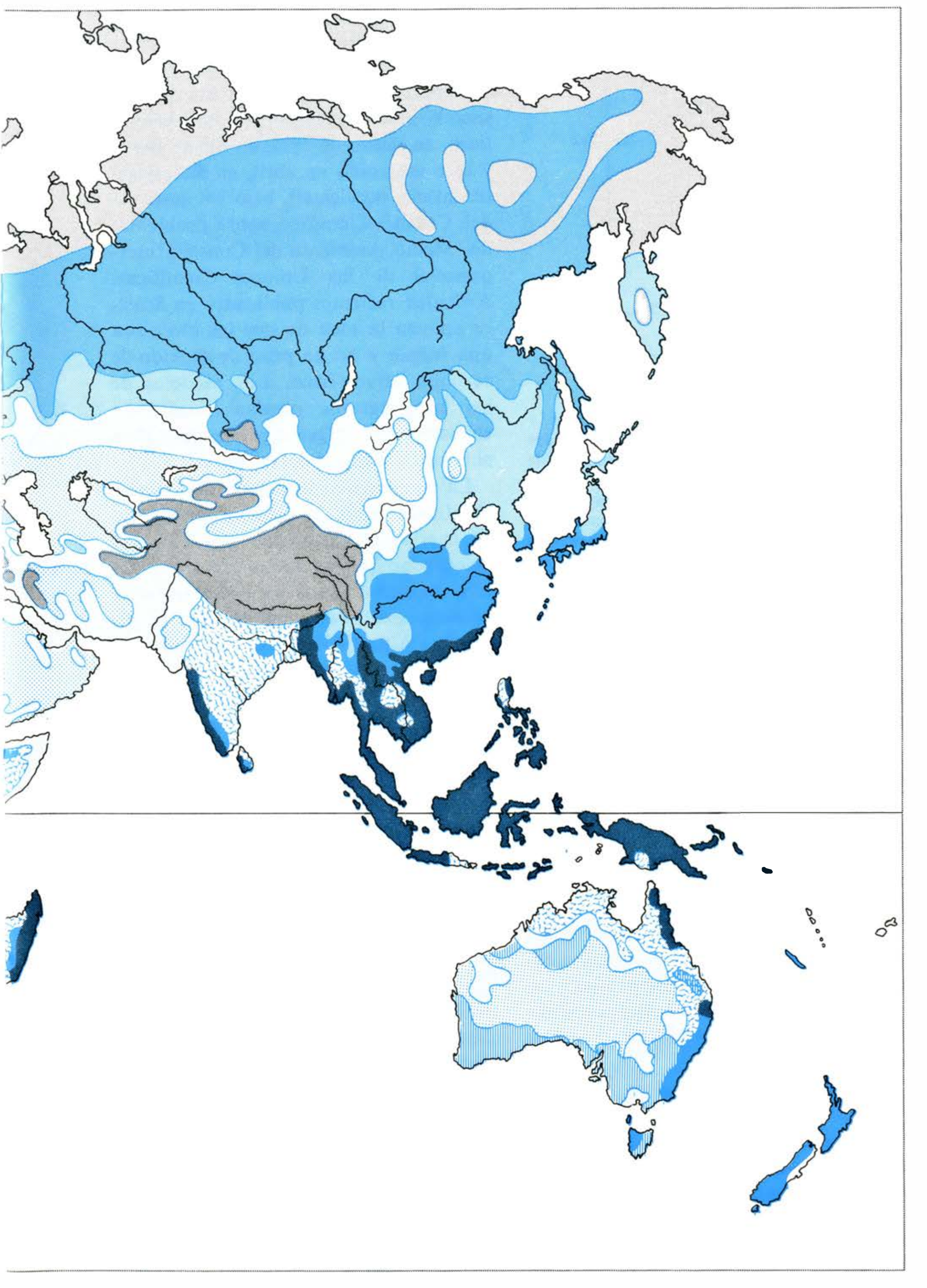
Cuando se intenta trazar un mapa de flujos que muestre el transporte neto de carbono, desde un sistema a otro, se encuentra que las estimaciones obtenibles varían ampliamente en calidad. Los datos conocidos con mayor precisión son los de emisión de dióxido de carbono por la combustión a nivel mun-



DISTRIBUCION DEL CARBONO almacenado en las plantas (mapa mundial basado en el trabajo de H. Brockman Jerosch). La cantidad total de carbono contenido en la biomasa terrestre es de unos 830 × 10^{15} gramos. En comparación, el carbono de toda la biomasa oceánica es despreciable: menos de 2 × 10^{15} gramos. Un 40 por ciento del carbono de las plantas reside en los bosques de lluvias tropicales.

dial de combustibles fósiles, que suelen cifrarse alrededor de 5 × 10^{15} gramos de carbono por año y las de incremento del contenido en dióxido de carbono del aire, equivalente, aproximadamente, a 23 × 10^{15} gramos de carbono por año. Esto implica que 2,7 × 10^{15} gramos de carbono procedente de los combustibles fósiles han de ser captados por alguna combinación de procesos terrestres y oceánicos. Supongamos por el momento que los biota terrestres representen un caudal estable de carbono que ni aumente ni disminuya la cantidad de dióxido de carbono en la atmósfera. En este caso (que se verá improbable) el océano debería captar y retener 2,7 × 10^{15} gramos de carbono por año. ¿Hay pruebas que confirmen esa tasa de remoción?

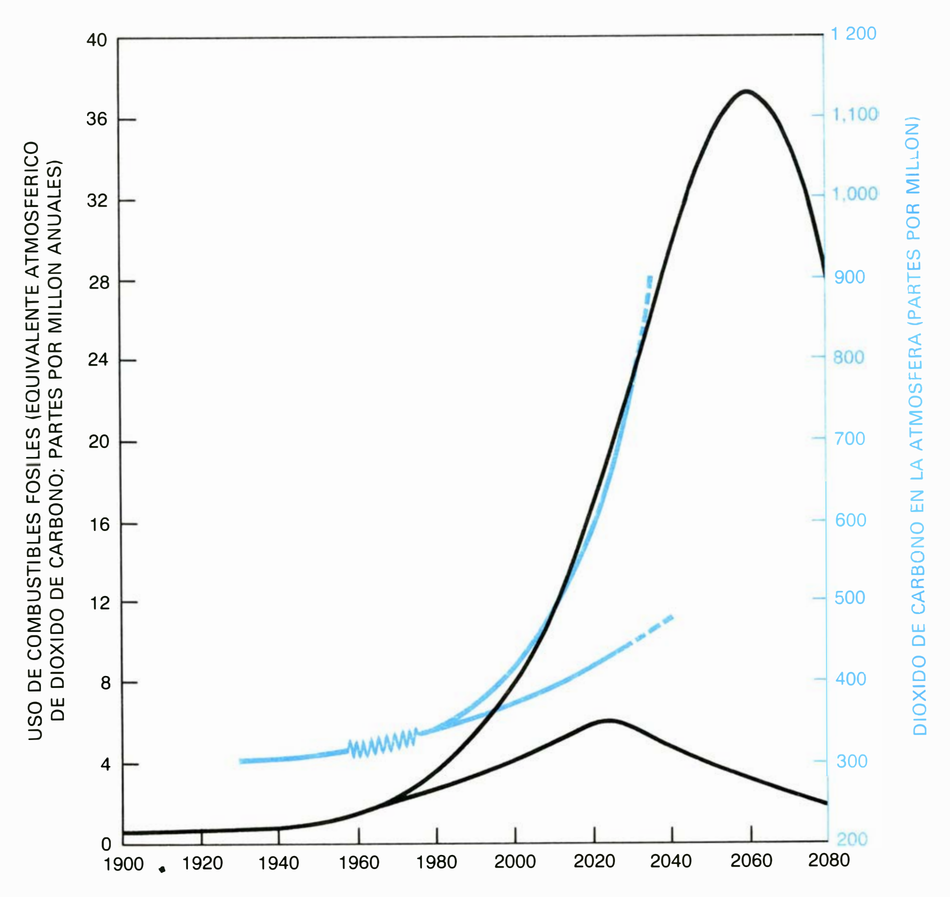
De acuerdo con las mejores estimaciones de los oceanógrafos químicos, se hace difícil explicar la forma en que el océano podría absorber anualmente la cantidad citada. Sus análisis están basados en un detallado conocimiento de la cantidad de carbono contenida en el sistema carbónico-carbonatos de la capa superficial, combinado con modelos cuidadosamente construidos de los mecanismos de mezcla en los océanos. Los isótopos radiactivos carbono 14 y tritio (hidrógeno 3), producidos en gran cantidad por las pruebas atómicas de la década de los años cincuenta y primeros años de la década de los sesenta, se han usado como elementos trazadores en el examen de las velocidades de intercambio entre el agua de la capa de mezcla y el agua de las profundidades oceánicas. Estos estudios parecen demostrar que la tasa de mezcla es realmente muy baja. La transferencia del carbono desde la atmósfera, a través de la capa superficial, hasta las aguas de la región abisal difícilmente excede de 2,5 × 10^{15} gramos por año, según algunos cálculos. En resumen, y de acuerdo con estos estudios, los océanos parecen ser escasamente adecuados para captar y retener la diferencia entre los 5 × 10^{15} gramos por año de carbono emitido por la combustión de combustibles fósiles y los 2,3 × 10^{15} gramos que la atmósfera retiene realmente.



Otro 14 por ciento permanece en los bosques tropicales estacionales. Los bosques de todas las latitudes contienen el 90 por ciento del carbono almacenado en los ecosistemas terrestres y marinos. La reducción de la superficie de los bosques ha elevado la concentración de dióxido de carbono.

Este equilibrio de flujos ha de reexaminarse completamente si el caudal biótico de carbono, en vez de aumentar o estar en equilibrio con el carbono atmósferico, es realmente una fuente de emisión de dióxido de carbono. Whittaker y Likens han tabulado recientemente la información existente acerca de la importancia de varios sectores de los biota. Este trabajo demuestra que los mayores almacenes de carbono de los biota son los bosques. También demuestra que la mayor parte de la fotosíntesis tiene lugar en tierra, y no en los océanos como se creía antes fundándose en antiguas suposiciones sobre la producción primaria neta de los océanos (producción primaria neta es la cantidad neta de carbono fijado, o materia orgánica, conseguida por la fotosíntesis después de que las necesidades respiratorias de la planta han sido cubiertas. Esta es la materia orgánica utilizable para el crecimiento de la planta y, por consiguiente, para almacenamiento o para uso por los animales y organismos en descomposición).

Quizás el hallazgo más significativo del estudio de Whittaker-Likens es el de que los bosques de los trópicos lluviosos, con sus grandes árboles, representan el mayor reservorio individual de carbono de los biota y aportan asimis-



PREVISIONES SOBRE EL CONSUMO DE COMBUSTIBLES FOSILES *(curvas negras)* **y contenido en dióxido de carbono de la atmósfera** *(color).* **Las gráficas muestran las tasas estimadas de aumento, máximas y mínimas, para ambos parámetros. Las tasas de consumo de combustible se han tomado de un estudio recientemente realizado en el Oak Ridge National Laboratory. La previsión mínima calcula un incremento anual del 2 por ciento hasta el año 2025, seguido de un descenso simétrico. La previsión máxima estima una tasa de crecimiento del 4,3 por ciento anual hasta que dicha tasa quede limitada por el agotamiento de las fuentes de combustible, a mediados del próximo siglo. La incertidumbre de tales predicciones hace que las apreciaciones sobre el contenido futuro en dióxido de carbono atmosférico sean extremadamente arriesgadas. La imprecisión viene agravada por la posibilidad, recientemente conocida, de que la destrucción de los bosques esté liberando grandes cantidades de dióxido de carbono a la atmósfera. La zona quebrada de la curva corresponde a las mediciones realizadas en Mauna Loa desde 1958. En la actualidad, el combustible fósil quemado cada año emite a la atmósfera una cantidad de dióxido de carbono equivalente a un incremento en la concentración de 2 partes por millón. Alrededor de la mitad de tal cantidad de dióxido de carbono es la que realmente queda retenida en la atmósfera terrestre.**

mo la mayor producción primaria neta total. Esta observación realza la importancia de los bosques tropicales en el balance terrestre del carbono. Si estos y otros bosques se ven sometidos a una tala rápida y se emite el carbono almacenado, tienen capacidad suficiente para contribuir de forma significativa al aumento de la cantidad de dióxido de carbono de la atmósfera. De forma inversa, la repoblación forestal de tierras absorbe cierta cantidad de dióxido de carbono, retardando la tasa de aumento de dicho gas.

Curiosamente, la cuestión de si la reserva biótica de carbono está creciendo o disminuyendo ha sido objeto de controversia tan sólo recientemente. A modo de ejemplo, en un ciclo de conferencias celebrado el año 1970, bajo el título "Estudio de los problemas críticos del medio ambiente", organizado por el Instituto de Tecnología de Massachusets y en Williamstown, Mass., se estableció la suposición de que la reserva biótica de carbono debía estar aumentando, y se alentaba así a los oceanógrafos a creer que sus modelos de circulación eran los adecuados. Otros ciclos subsiguientes de conferencias vinieron a reforzar esta suposición, aunque en 1972 surgieron ciertas dudas en un seminario titulado "Carbono y biosfera", celebrado en el Brookhaven National Laboratory.

El primer reto serio al modelo propuesto hasta entonces por los oceanógrafos se lanzó en dos comunicaciones presentadas en la Reunión Dahlem sobre biogeoquímica, celebrada en Berlín occidental en noviembre de 1976. En una de las comunicaciones, R. A. Houghton y yo estimamos que los biota podrían estar emitiendo anualmente a la atmósfera tanto carbono como la combustión de los combustibles fósiles. En la otra comunicación, Bolin llegó a un valor más bajo de emisión de carbono de procedencia biótica, teniendo en cuenta los datos de talas de bosques obtenidos por la FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación). La estimación de Bolin fue de, aproximadamente, 10^{15} gramos de carbono. Este tema dominó la Reunión Dahlem originando muchas discusiones.

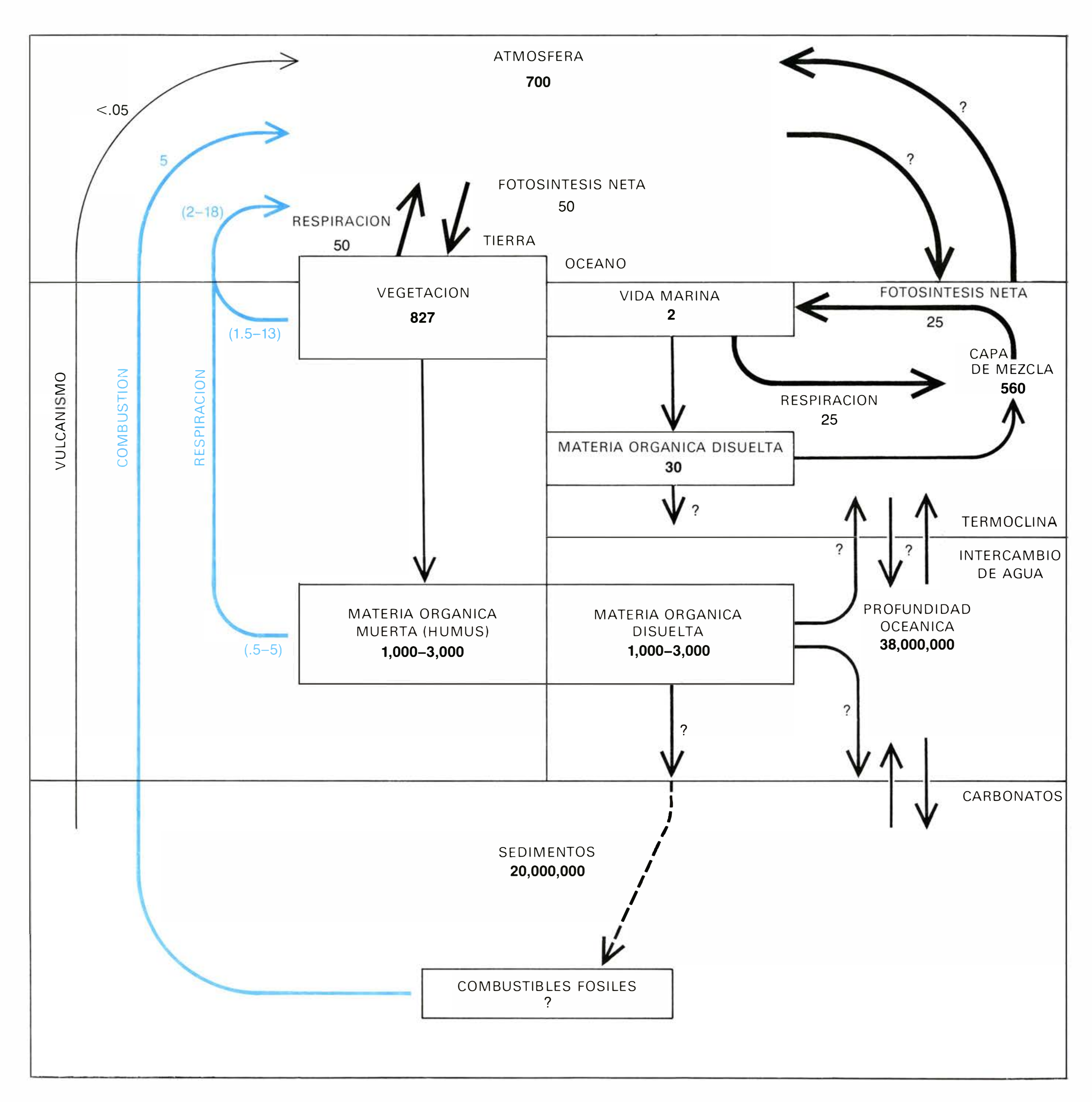
La posibilidad de que exista una liberación significativa de dióxido de carbono a partir de los biota también dominó en dos ciclos de conferencias subsiguientes, uno organizado por el Department of Energy (antes Energy Research and Development Administration), en marzo de 1977, en Miami Beach y otro sostenido en abril, en Ratzeburg (Alemania occidental), bajo los auspicios del Comité Científico sobre problemas del Medio Ambiente del Consejo Internacional de las Uniones Científicas. Artículos recientes publicados en *Science* apoyan la idea de que los biota son una fuente y no un pozo de dióxido de carbono atmosférico. Los modelos de los oceanógrafos quedan claramente en entredicho. ¿Qué significa esto para el balance mundial del carbono?

La respuesta dista mucho de hallarse clarificada. En la atmósfera están entrando cantidades sustancialmente mayores de carbono que las que se almacenan en ella. Además de los 5×10^{15} gramos de carbono emitidos anualmente a la atmósfera a partir de los combustibles fósiles, de 4×10^{15} gramos a 8×10^{15}, y quizá más, pueden estar siendo reliberados de forma corriente a través de la destrucción de bosques y de la acelerada oxidación del humus. De esta cantidad combinada (de 9×10^{15} a 13×10^{15} gramos de carbono por año), solamente $2{,}3 \times 10^{15}$ gramos se acumulan en la atmósfera. Los restantes 7×10^{15} a 11×10^{15} gramos, y posiblemente más, se almacenan en alguna parte de la tierra. Pero, ¿dónde? Como hemos visto, según los actuales modelos, los océanos están capacitados para captar menos de 3×10^{15} gramos de carbono por año. Los oceanógrafos están revisando sus hipótesis para comprobar si no han pasado por alto mecanismos capaces de retener cantidades adicionales de carbono.

Debido a la importancia del tema, los familiarizados con los ecosistemas terrestres están examinando cuidadosamente sus propios datos, especialmente los relacionados con cambios en las masas forestales. ¿En qué grado podemos estar seguros de que las comunidades terrestres son una fuente y no un pozo? El

problema se ha orientado en varias direcciones en los últimos meses. Los datos no son todo lo buenos que cabría desear, pero parecen impulsar a los familiarizados con ellos.

El análisis está basado, en primer lugar, en el conocimiento de las magnitudes relativas de las cantidades de carbono mantenidas dentro de los biota y en la producción primaria neta de cada una de las más importantes comunidades vegetales de la tierra. El estudio de Whittaker-Likens muestra, por ejemplo, que los bosques de lluvia tropical contienen alrededor del 42 por ciento de todo el carbono de la vegetación terrestre y son responsables del 32 por ciento de la producción primaria total neta. Las diferentes superficies forestales (tropicales, templadas y boreales) contienen el 90 por ciento de todo el carbono retenido en la vegetación y contribuyen en más de un 60 por ciento en la producción neta primaria. Otro contribuyente singular de



EL EQUILIBRIO GLOBAL del carbono reseña sus más importantes depositarios, así como las tasas de intercambio entre los diversos reservorios en contacto. Las cantidades vienen expresadas en unidades de 10^{15} gramos, o miles de millones de toneladas métricas. Las emisiones anuales a la atmósfera, producidas por actividades humanas, se muestran en color. Las plantas terrestres tienen una fijación neta de alrededor de 50×10^{15} gramos de carbono. Este carbono es consumido y rápidamente respirado por varios organismos terrestres o bien queda almacenado en la masa de las plantas. El equilibrio entre fijación (fotosíntesis neta) y almacenamiento más respiración total de todos los organismos terrestres determina si existe un flujo neto de carbono desde, o hacia, el conjunto de seres vivos. Muchos biólogos creen hoy que ha existido desde hace tiempo un flujo neto de dióxido de carbono, desde los biota hacia la atmósfera, y que esta tendencia continúa. El carbono fijado por los organismos marinos es respirado o almacenado. Comúnmente se ha aceptado que la mayor parte de este carbono es respirado inmediatamente y reciclado. Hoy en día, parece posible que el transporte de materia fecal desde la superficie hacia las profundidades oceánicas haya transportado hasta las mismas más carbono del que se creía. Esta transferencia supliría a la difusión, normalmente lenta, del dióxido de carbono en las capas superficiales de los océanos, donde entra en equilibrio con el sistema carbónico-carbonatos. Aunque las profundidades oceánicas son un pozo prácticamente ilimitado para el dióxido de carbono, el gas debe llegar a la capa de mezcla y entonces penetrar en la termoclina.

importancia en esta producción neta primaria son las sabanas o tierras de pastos que aportan un 12 por ciento de la misma y representan solamente alrededor del 3 por ciento de la masa de carbono. Todos los terrenos cultivados de la tierra presentan, aproximadamente, un 8 por ciento de la producción primaria total neta y menos de un 1 por ciento del contenido de carbono. La estimación de Whittaker-Likens se acerca a un valor intermedio entre los diversos resultados obtenidos en los análisis realizados bajo la dirección de P. Duvigneaud, de la Universidad de Bruselas, en la reunión de Ratzeburg. Puesto que los diversos estudios confirman la importancia de los bosques, y particularmente de las selvas tropicales, es esencial establecer si estos ecosistemas están modificando su extensión o no, y si lo están haciendo, a qué velocidad.

Los datos escasean. Disponemos de las experiencias de Henry C. Darby, geógrafo británico que, en 1954, publicó una evaluación de los cambios en la vegetación europea desde el año 900 después de Cristo hasta el año 1900. En dicho período, se redujo la cobertura forestal de Europa occidental, pasando de un 90 por ciento a un 20 por ciento. Un cambio similar tuvo lugar de forma más temprana en las tierras mediterráneas, en especial en la región de levante. Esta gran reducción de bosques liberó una cantidad tal de carbono que representó una fracción significativa de la cantidad total de dicho elemento que antes contenía la atmósfera. Es razonable suponer que desde 1900 la industrialización continuada y el crecimiento de población se han traducido en cambios similares en los bosques de otros lugares.

Timothy Wood y Daniel B. Botkin, del Ecosystem Center de Woods Hole, llevaron a cabo recientemente un estudio sobre los cambios producidos en el área forestal de Nueva Inglaterra, desde la llegada de los colonizadores europeos. Dichos autores encontraron que existió una incesante reducción de los bosques hasta el año 1900, si bien desde entonces se ha registrado un período de recuperación a causa de la recesión de la agricultura y de la expansión de los bosques sobre las primitivas tierras agrícolas. La recuperación no ha supuesto, sin embargo, un aumento de la reserva de carbono que fuera equivalente a la de los bosques primitivos. Los bosques se han talado con regularidad y nunca han alcanzado ni la altura ni la extensión de los originales; la reserva actual de carbono no es más que la mitad de la original. Los datos más recientes indican que se ha detenido el aumento del área forestal probablemente a causa de la renovada expansión de la agricultura y a la intensificación de la tala de árboles.

El estudio de Wood y Botkin muestra que en la Zona Templada, en un bosque en proceso de recuperación, el carbono almacenado aumentó en un 3 o 4 por ciento anual de la producción primaria neta, durante todo el período de recuperación (unos 70 años). Si se almacenase una fracción similar de la producción primaria neta de otros bosques de la Zona Templada, y se retuviera en el humus una igual cantidad adicional, la cantidad de carbono acumulado en todos los bosques de la Zona Templada de la tierra se elevaría, aproximadamente, a 0,5 × 10^{15} gramos por año. La experiencia de los bosques de Nueva Inglaterra sugiere que la repoblación forestal de las zonas templadas no se muestra hoy probablemente como un gran reservorio del dióxido de carbono atmosférico.

Al propio tiempo ha existido simultáneamente una continua expansión de la

	SUPERFICIE (10^6 KILOMETROS CUADRADOS)	PRODUCCION PRIMARIA NETA (10^{15} GRAMOS DE CARBONO POR AÑO)	BIOMASA VEGETAL (10^{15} GRAMOS DE CARBONO)
BOSQUES DE LLUVIAS TROPICALES	17.0	16.8	344.0
BOSQUES TROPICALES ESTACIONALES	7.5	5.4	117.0
BOSQUES PERENNES (ZONA TEMPLADA)	5.0	2.9	79.0
BOSQUES HOJA CADUCA (ZONA TEMPL.)	7.0	3.8	95.0
BOSQUES BOREALES	12.0	4.3	108.0
BOSQUES CLAROS Y MATORRALES	8.5	2.7	22.0
SABANAS	15.0	6.1	27.0
TIERRAS DE PASTO (ZONA TEMPLADA)	9.0	2.4	6.3
TUNDRAS Y PRADERAS ALPINAS	8.0	0.5	2.3
DESIERTO DE MONTE BAJO	18.0	0.7	5.9
ROCAS, HIELO Y ARENA	24.0	0.03	0.2
TIERRAS CULTIVADAS	14.0	4.1	6.3
MARISMAS Y PANTANOS	2.0	2.7	13.5
RIOS Y LAGOS	2.0	0.4	0.02
TOTAL CONTINENTAL	149.0	52.8	826.5
OCEANO ABIERTO	332.0	18.7	0.45
ZONAS DE AFLORAMIENTO	0.4	0.1	0.004
PLATAFORMA CONTINENTAL	26.6	4.3	0.12
LECHO DE ALGAS Y ARRECIFES	0.6	0.7	0.54
ESTUARIOS	1.4	1.0	0.63
TOTAL MARINO	361.0	24.8	1.74
TOTAL MUNDIAL	510.0	77.6	828.0

LAS PRINCIPALES COMUNIDADES DE PLANTAS de la tierra vienen reseñadas en esta tabla, junto a la superficie que ocupan, su producción primaria neta y la cantidad de carbono que contienen. Producción primaria neta es la cantidad de carbono que una comunidad vegetal produce anualmente para cosecha o para el soporte de varios organismos de consumo, ya sean salvajes o domesticados por el hombre. Aunque tan sólo un 30 por ciento de la superficie del planeta está cubierta por tierra, la producción primaria neta de la vegetación terrestre es algo mayor que el doble de la producción primaria de los océanos. La cantidad de carbono almacenado en las plantas terrestres es unas 500 veces mayor que la cantidad contenida en los ecosistemas marinos. El contenido en carbono de los árboles es aproximadamente igual al carbono de la atmósfera. Los valores de la tabla fueron compilados recientemente por R. H. Whittaker y Gene E. Likens.

agricultura sobre otras regiones forestales, una continua tala de bosques en otros lugares y una intoxicación general de la tierra a consecuencia de las actividades humanas. Los mayores daños se han sufrido en los bosques tropicales. La superficie forestal de mayor extensión actualmente existente es la Cuenca Amazónica; nosotros hemos solicitado datos sobre la tasa de tala de dicha Cuenca Amazónica. No se dispone de memoria alguna capaz de aplicarse satisfactoriamente a toda la Cuenca Amazónica sin incurrir en los riesgos asociados con toda extrapolación. A pesar de todo, Lawrence S. Hamilton, de la Cornell University, y colaboradores han publicado recientemente una serie de informes sobre los cambios en los bosques venezolanos.

Uno de los informes, debido a J. P. Veillon, señala que el área forestal de los llanos occidentales de Venezuela se ha reducido en un 33 por ciento entre 1950 y 1975. Hamilton, en una introducción a la serie, cita los datos de la FAO que "sugieren una estimación de que los bosques húmedos son clareados anualmente en un 0,6 a 1,5 por ciento del área aún existente". Diversos estudios sobre otros sectores de la Cuenca Amazónica testimonian la alta tasa de desarrollo de autopistas, el avance de la agricultura a expensas de los bosques y la disminución de la revegetación cuando se clarean los bosques. Prácticamente, ningún investigador lúcido con experiencia en bosques tropicales cree que la Cuenca Amazónica estará a salvo de importantes destrozos en los próximos 30 años, a consecuencia de la tala y el clareado de los bosques.

Las mejores estimaciones que he podido obtener suponen que la tasa de clareado de los bosques tropicales es probablemente del orden del 0,5 al 1,5 por ciento anual del área existente. Si la tasa es del 1 por ciento y si la mayor parte de dicho clareado se produce a causa de la penetración de la agricultura y de las tierras de pasto en antiguas zonas forestales, la liberación de carbono, a partir tan sólo de los boques, sería de unos $4,5 \times 10^{15}$ gramos por año. Aplicando estos cálculos al conjunto de todos los biota terrestres, teniendo en cuenta el almacenamiento adicional producido por la revegetación, mis colegas y yo hemos calculado recientemente que la emisión neta de dióxido de carbono a la atmósfera, procedente de los biota, es de 6×10^{15} gramos de carbono por año. Es difícil calcular la magnitud de una emisión adicional producida por descomposición del humus del suelo; nosotros hemos estimado esta pérdida en 2×10^{15} gramos por año. La cantidad total liberada será, por tanto, de unos 8×10^{15} gramos por año. No obstante, las imprecisiones son tan grandes, que hemos supuesto que la pérdida real a partir de los seres vivos podría oscilar entre 2×10^{15} y 18×10^{15} gramos por año. Aunque no hay que olvidar las limitaciones de los datos sobre los que se han hecho estas suposiciones, resulta difícil evitar la conclusión de que la destrucción de los bosques de la tierra está añadiendo dióxido de carbono a la atmósfera a un ritmo comparable al aumento que procede de la combustión de combustibles fósiles, y si se incluye la oxidación del humus, a un ritmo apreciablemente mayor.

BOSQUES	BIOMASA VEGETAL (10^3 GRAMOS DE CARBONO POR METRO CUADRADO)	PRODUCCION PRIMARIA NETA (GRAMOS DE CARBONO POR METRO CUADRADO Y AÑO) RANGO	MEDIA
TROPICALES (HUMEDOS)	3.0–36.0	450–1,600	990
TEMPLADOS	3.0–90.0	270–1,125	560
BOREAL	3.0–18.0	180–900	360
SABANAS	0.1–7.0	90–900	400
TIERRAS DE PASTO	0.1–2.3	90–675	270
TIERRAS CULTIVADAS	0.2–5.4	45–2,800	290

TRANSFORMACION DE TIERRAS SALVAJES en terrenos agrícolas. Suele traducirse en una reducción aguda del carbono almacenado en la biomasa y en una reducción algo menor del dióxido de carbono separado anualmente de la atmósfera y fijado por fotosíntesis. Esta tabla, basada también en el trabajo de Whittaker y Likens, recoge los valores de las evaluaciones publicadas sobre la biomasa permanente y la producción primaria neta de los principales ecosistemas naturales.

Estas observaciones han sorprendido a numerosos agricultores y silvicultores. La creencia general era de que las modernas prácticas agrícolas y forestales aumentaban la producción neta primaria de los sistemas terrestres bajo control humano y que la mejora de dicha producción neta primaria debía haberse traducido en un almacenamiento más rápido en el humus de dichos suelos sometidos a control humano que no en los suelos naturales. Pero la productividad promedio de la agricultura, medida en las unidades usadas por los ecólogos (total de materia orgánica utilizable por el hombre u otros animales), es sustancialmente menor que la producción primaria promedio de las comunidades naturales que la agricultura ha reemplazado. Las comunidades de plantas agrícolas no se cultivan para almacenar grandes cantidades de carbono, sino para el rápido reciclado de las reservas de carbono a través de su utilización por el hombre. Esto significa que las tierras agrícolas no almacenan tanto carbono como los bosques reemplazados por ellas. De forma similar, las tierras de pasto dedicadas a la agricultura tienden a perder materia orgánica del suelo por desintegración y a no acumular materia orgánica adicional.

Los bosques sometidos a un cultivo y explotación intensos, aunque pueden aumentar su rendimiento en madera, tienden a disponer, en el tiempo de tala, de una reserva menor que los bosques silvestres a los que han reemplazado. Esto significa de nuevo que se ha reducido el tiempo de reposición y que la reserva de carbono nunca alcanza el mismo tamaño que la de los bosques originales. De esta forma, la substitución de bosques primarios por bosques secundarios cultivados y explotados con vistas a la obtención de madera para la construcción o de pulpa se traduce en un aumento de la emisión neta de dióxido de carbono. Ocurre una nueva emisión de carbono a partir de la descomposición del humus. Esta descomposición es estimulada por la tala de los bosques y puede prolongarse durante las primeras etapas de recuperación que siguen a dicha tala. Todos estos factores conducen a un aumento neto de la emisión de carbono cuando los bosques primarios ceden su lugar a los explotados por la mano del hombre.

Pruebas de otro orden parecen apoyar la idea de que los biota terrestres constituyeron, durante numerosas décadas, una fuente de emisión de dióxido de carbono y no un pozo. Comparando las proporciones de isótopos de carbono en árboles, Minze Stuiver, de la Washington University, calculó que en el siglo transcurrido entre 1850 y 1950 los biota pudieron haber liberado anualmente a la atmósfera $1,2 \times 10^{15}$ gramos de carbono.

Durante el mismo período, la emisión a partir de combustibles fósiles representó en promedio $0,6 \times 10^{15}$ gramos de carbono por año.

Stuiver utilizó la relación en que se encuentran el carbono 12 y el carbono 13 en la atmósfera, en los biota y en los combustibles fósiles. Los seres vivos y los combustibles fósiles están ligeramente enriquecidos en el isótopo más ligero, el carbono 12. Además, el tercer isótopo, el carbono 14, se halla presente tan sólo en la atmósfera y en los biota. El carbono 14 se produce en las capas altas de la atmósfera a consecuencia del bombardeo por rayos cósmicos del isótopo común del nitrógeno, el nitrógeno 14. También se ha producido en grandes cantidades en las pruebas atómicas atmosféricas. Ya que el carbono 14 tiene una vida media de unos 6000 años, ha desaparecido hace tiempo de los combustibles fósiles, los cuales se formaron hace millones de años. Por tanto, la combustión de los combustibles fósiles emite unas cantidades de carbono pobres en el isótopo 14 y tiende a diluir el carbono 14 existente en la atmósfera. Midiendo la concentración de los diversos isótopos de carbono en los anillos de árboles de edad conocida, y comparando las proporciones existentes con las que registramos en la atmósfera y en los combustibles fósiles, Stuiver pudo valorar la cantidad de carbono liberado por los biota. Las mediciones, desde el punto de vista técnico, son difíciles; por otra parte, las proporciones entre los isótopos del carbono en los árboles vienen afectadas por varios factores del medio ambiente que hacen que los resultados sean mucho menos definitivos de lo deseable. En todo caso, la técnica ofrece un importante medio adicional de averiguar la emisión de dióxido de carbono por los biota.

A la vista de las pruebas existentes, parece estar fuera de toda duda que el aumento del contenido en dióxido de carbono atmosférico no sólo se debe a la combustión de los combustibles fósiles, sino también a la destrucción de los bosques. G. Evelyn Hutchinson, de la Yale University, estableció este punto en un capítulo del libro "La tierra como planeta" *(The Earth as a Planet)* editado por Gerard P. Kuiper en 1954. Hutchinson supuso que la liberación de dióxido de carbono por los seres vivos era aproximadamente igual a la causada por combustión de los combustibles fósiles. Las mejores pruebas de que hoy disponemos indican que esta relación persiste.

Hay suficiente carbono soportado por los biota para que la relación continúe durante una década o dos a través del esperado pico en el consumo mundial de petróleo. En este tiempo, si no antes, la tierra habrá sufrido un cambio climático como consecuencia de la acumulación de dióxido de carbono en la atmósfera, si es que realmente tal cambio tiene lugar. El si tendrá o no tendrá lugar depende de la importancia del efecto causado por el dióxido de carbono, efecto que es extremadamente difícil de tabular con certeza. Se reconoce hoy que los cambios en la emisión de energía solar, en el grado de reflectividad de la tierra (bajo la influencia del recubrimiento por nubes, nieve y hielo) y otros factores, condicionan el clima. Si el contenido en dióxido de carbono atmosférico es suficientemente grande o no, para constituirse en factor determinante de un cambio climático, es un hecho por comprobar. Si el efecto del dióxido de carbono es el dominante, la tierra se calentará de forma diferenciada, con temperaturas crecientes en dirección a los polos. Cabe esperar que tal cambio desplazaría las zonas desérticas hacia los polos, aumentando las zonas áridas y reduciendo las utilizables para la agricultura. La perspectiva no es esperanzadora para un mundo cuya población puede doblarse en los próximos 30 o 35 años.

Si existiese una prueba tajante en pro de un riesgo de cambio climático insoportablemente grande en las próximas décadas, las medidas a tomar resultarían inequívocamente claras. Habría que reducir la combustión de los combustibles fósiles con el objeto de disminuir esta fuente de dióxido de carbono. Habría que dedicar numerosos esfuerzos para prevenir la tala de los bosques primarios en todo el mundo, para aumentar las zonas forestales y permitir que en tales zonas se desarrollasen masivamente los árboles. El que tales medidas se llevaran a cabo resulta más que dudoso; los problemas sociales que derivarían de su aplicación no serían, obviamente, menos graves.

Se han hecho otras recomendaciones; una de ellas, basada en que la disponibilidad de fósforo es considerada por ciertos autores como factor limitante de la producción neta primaria de los océanos, propone que las naciones más desarrolladas deberían aplicar parte de su potencial industrial a la obtención de fósforo con el objeto de transferirlo tan rápidamente como fuese posible a los océanos relativamente infértiles, estimulando de este modo la fotosíntesis y el almacenamiento de carbono. Este esquema es atractivo, ya que parece ofrecer una forma de acelerar el almacenamiento de carbono en los océanos. Existe, sin embargo, la posibilidad de que un aumento adicional de la fotosíntesis en el océano vendría compensado por un aumento equivalente de la respiración, de forma que no se produciría un aumento neto del carbono almacenado. Existen numerosas dudas acerca de la posibilidad de realización de este programa, incluyendo la de si realmente el fósforo es un factor limitante en las aguas oceánicas. Existe la razonable probabilidad de que el nitrógeno también desempeñe un papel limitante, en cuyo caso la estimulación de la fotosíntesis marina resultaría mucho más ardua de lo que se creía.

En todo caso, las experiencias recientes nos han recalcado a todos los que trabajamos en estos campos que aspectos importantes del balance mundial del carbono permanecen hoy, en líneas generales, desconocidos. Según parece, se puede mejorar el marco teórico de alguno de estos aspectos notables gracias a las mediciones realizadas sobre los cambios de extensión y estructura de los bosques de todo el mundo, mediante fotografías obtenidas con satélites artificiales. Asimismo, se están desarrollando investigaciones para comprobar si los mecanismos bióticos podrían facilitar la transferencia a las profundidades marinas de una cantidad considerablemente mayor de carbono de la que se supone se está transfiriendo en la actualidad. No existen, sin embargo, perspectivas inmediatas de poder determinar de forma altamente precisa todos los detalles del balance mundial del carbono.

Los riesgos potenciales asociados a un rápido incremento del contenido en dióxido de carbono atmosférico surgirán ampliamente en las próximas décadas e influirán poderosamente en algunas decisiones, verbigracia: aceleración del desarrollo de las plantas nucleares de energía eléctrica en detrimento de las centrales térmicas basadas en la combustión de carbón; y preservación de las zonas forestales en lugar de restringirlas (y si se han de preservar las zonas forestales, buscar la forma de conseguir las nuevas tierras que sin duda será necesario dedicar a la agricultura). Prácticamente, no existe ningún aspecto de la política nacional e internacional que se muestre inconmovible ante la perspectiva de un cambio climático a escala mundial. El dióxido de carbono, hasta ahora un gas aparentemente inocuo, presente a nivel de trazas en la atmósfera, puede pasar muy pronto a desempeñar un papel central al constituir una amenaza, nada negligible, del orden mundial.

La sustitución quirúrgica de la articulación de la rodilla

Los progresos hechos en materiales, diseños y técnicas quirúrgicas han mejorado tanto los resultados de las prótesis de rodilla que pueden aportar importantes beneficios a los pacientes afectos de artritis grave

David A. Sonstegard, Larry S. Matthews y Herbert Kaufer

La carga que soporta la articulación de la rodilla humana al realizar actividades físicas habituales, como pasear, correr, arrodillarse, subir escaleras y sentarse o levantarse de un asiento, puede superar en cinco veces la totalidad del peso corporal. Hoy en día, un gran número de personas presentan a sus rodillas otras (y a menudo desaforadas) exigencias al participar en ciertos deportes (fútbol, rugby, tenis y atletismo). No es de extrañar que mucha gente llegue a la vejez con una o quizás ambas rodillas tan deterioradas como para quedarse inválidos.

Se puede ahora sustituir dichas rodillas por un dispositivo mecánico que imita los complejos y delicados movimientos de que es capaz una rodilla sana. Actualmente las prótesis se aplican sólo a pacientes con alto grado de invalidez y que probablemente no exigirán grandes esfuerzos a esa articulación. La rodilla artificial no constituye todavía una solución para personas en plena y juvenil actividad, primordialmente porque las modernas ténicas de fijación de la prótesis al hueso son incapaces de soportar las cargas exigidas por tales individuos. En este trabajo describiremos diversos modelos de prótesis, insistiendo especialmente en la rodilla esferocéntrica, creada por nuestro grupo de la Universidad de Michigan.

Las investigaciones sobre prótesis de rodilla han surgido a raíz del satisfactorio desarrollo de las sustituciones de cadera, a lo largo de los últimos 15 años. Gran parte del mérito en cuanto a la elección de materiales y configuraciones que servirían adecuadamente para la articulación coxofemoral, así como a las técnicas quirúrgicas para su implantación, corresponde al cirujano británico John Charnley, del Centro de Cirugía de la Cadera del Wrightington Hospital de Lancashire.

La articulación normal de la cadera es un mecanismo convexo-cóncavo que permite la realización de movimientos en todos los planos alrededor de su centro de rotación (la cadera artificial es un modelo semejante que proporciona el mismo grado de movilidad). Sin embargo, la articulación de la cadera está rodeada y protegida por una gran masa muscular y de tejido conjuntivo. La rodilla, por el contrario, tiene una gama de movimientos más limitada y está relativamente sin proteger por tejidos envolventes. Es mucho más vulnerable a los traumatismos directos, frenazos bruscos y torsiones. Por estas razones no ha resultado fácil conseguir con las prótesis de rodilla el éxito obtenido con las prótesis de cadera. Ni siquiera estas últimas pueden usarse en articulaciones gravemente traumatizadas, salvo en casos muy especiales. Las fracturas y las alteraciones anatómicas de importancia hacen imposible una alineación justa y un correcto soporte anatómico para la prótesis. Por ello, sólo los pacientes con artritis crónica grave han sido sometidos a sustituciones de cadera y rodilla.

En sentido literal, artritis significa "inflamación articular". El término se aplica a casi un centenar de afecciones, cuya sintomatología es: dolor persistente o recidivante, rigidez, laxitud o deformidad en una o varias articulaciones. La Arthritis Foundation ha estimado la incidencia de esta enfermedad en 363 millones de personas, que comprende el 10 por ciento de la población mundial. De ellos hay más de 50 millones en Estados Unidos, 20 millones de los cuales requieren asistencia médica y 3,5 millones están significativamente incapacitados.

Se desconoce la causa de la artritis. En el momento actual, no existe tratamiento médico efectivo para la mayoría de las variedades de esta enfermedad. Sus formas principales son la artritis reumatoide y la osteoartritis. Esta última es una forma degenerativa de artritis por desgaste (designada como "artrosis" en nuestro ambiente), que progresa con la actividad de la articulación y con la edad. La artritis reumatoide, enfermedad inflamatoria de los tejidos blandos, constituye la variedad más deformante de la enfermedad y, en consecuencia, la que causa mayor incapacidad funcional. El tratamiento de ambos tipos de artritis comprende fármacos, terapéutica física, reposo, vendajes y aplicación de calor. Pese a tales empeños, ambos tipos de artritis (osteoartritis y artritis reumatoide) pueden progresar hasta ocasionar un dolor intenso, inestabilidad articular, deformidad, limitación de la actividad e invalidez. En tal caso es probable la necesidad de acudir a la cirugía. La mayoría de prótesis de cadera y rodilla son consecuencia de estas indicaciones.

Para entender mejor lo que supone la sustitución de la articulación de la rodilla debe estudiarse la anatomía de la rodilla normal *(véase la ilustración de la página 31)*. La rodilla está compuesta por dos articulaciones o partes móviles, que transmiten la fuerza y el movimiento relativo. Una es la superficie patelofemoral, entre la rótula (el cerrojo de la rodilla) y el fémur (el hueso del muslo). La otra articulación es la tibiofemoral, entre la tibia (el más grueso de los huesos

largos situados entre la rodilla y el tobillo) y el fémur.

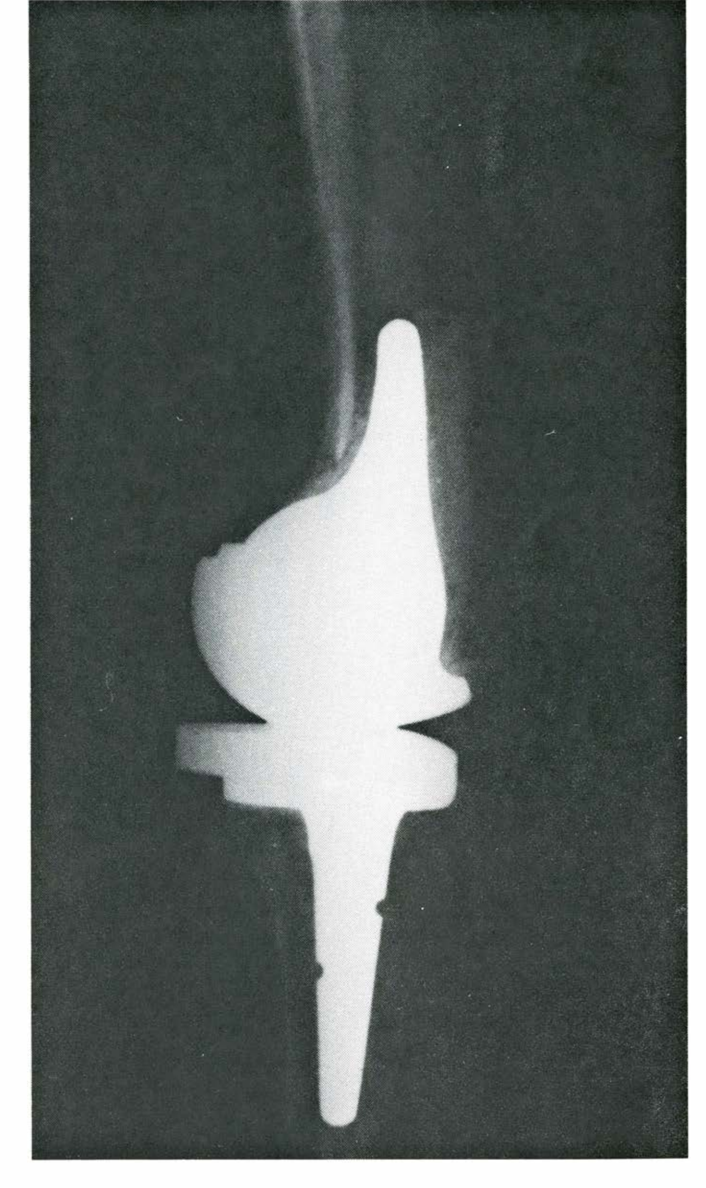
Hay tres grupos musculares básicos, sujetos al control del sistema nervioso, que suministran la fuerza o energía necesaria para el funcionamiento de la rodilla. El grupo muscular bíceps y el grupo muscular posterior entran en contracción para permitir que la pierna se doble por la rodilla. El grupo muscular cuádriceps, que aprovecha la ventaja mecánica de la rótula, endereza la pierna. Ambos movimientos reciben los nombres de flexión y extensión, respectivamente.

Los ligamentos de la rodilla, constituidos fundamentalmente por fibras colágenas paralelas, poseen una flexibilidad que permite una gran libertad de movimiento, aunque notablemente controlada. En toda su longitud, dichos ligamentos constituyen firmes conexiones interóseas relativamente inextensibles; lo cual proporciona estabilidad a la articulación y limitación de su motilidad. El ligamento rotuliano, al unir la rótula a la tibia, posibilita un movimiento de deslizamiento en la articulación rotulofemoral durante la flexión y la extensión. Otros dos pares de ligamentos, los laterales y los cruzados, estabilizan la articulación tibiofemoral.

La articulación tibiofemoral está provista de dos estructuras fibrocartilaginosas en forma de media luna: el menisco medio (situado en la parte interna de la rodilla) y el menisco lateral (en la externa). La cara superior de ambos meniscos es cóncava, con objeto de aumentar la profundidad de las mesetas tibiales que han de articularse con los cóndilos femorales, estructuras nudosas del extremo inferior del fémur. Las conexiones entre la superficie tibial, relativamente plana, y los meniscos permiten cierta movilidad de los meniscos respecto de la tibia. (Los frecuentes traumatismos deportivos catalogados como "desgarro del cartílago" afectan a un menisco o a ambos.)

Los cóndilos femorales lateral y medio son superficies complejas con radios de curvatura continuamente variables. Dichas superficies articulares, conjuntamente con los ligamentos y la forma geométrica de los meniscos y de la tibia, proporcionan la gama estabilizada y flexible de movimientos propia de la rodilla.





RODILLA ARTIFICIAL implantada en la pierna izquierda de un paciente afecto de artritis grave, vista en dos radiografías. Se trata de la prótesis esferocéntrica proyectada por los autores en la Universidad de Michigan, y que aparece en la portada de este número. Entre las ventajas distintivas del dispositivo esferocéntrico está el permitir casi los mismos movimientos triaxiales de que es capaz una rodilla normal y una deceleración gradual del movimiento, sin paradas abruptas. En dicho centro universitario se ha implantado más de 120 de estas prótesis.

El carácter del movimiento es específico. Por ejemplo, en la flexión y la extensión, la rodilla no actúa como un simple gozne, sino más bien como una sucesión infinita de centros de rotación de emplazamientos singular, cada uno de los cuales actúa en determinada orientación del fémur y de la tibia. Por tanto, la articulación tibiofemoral posee una combinación sutilmente concertada de movimientos de basculamiento, deslizamiento y rotación. La posición de extensión total, sin rotación alguna de la tibia respecto al fémur, constituye un mecanismo de cerrojo que permite estar cómodamente de pie.

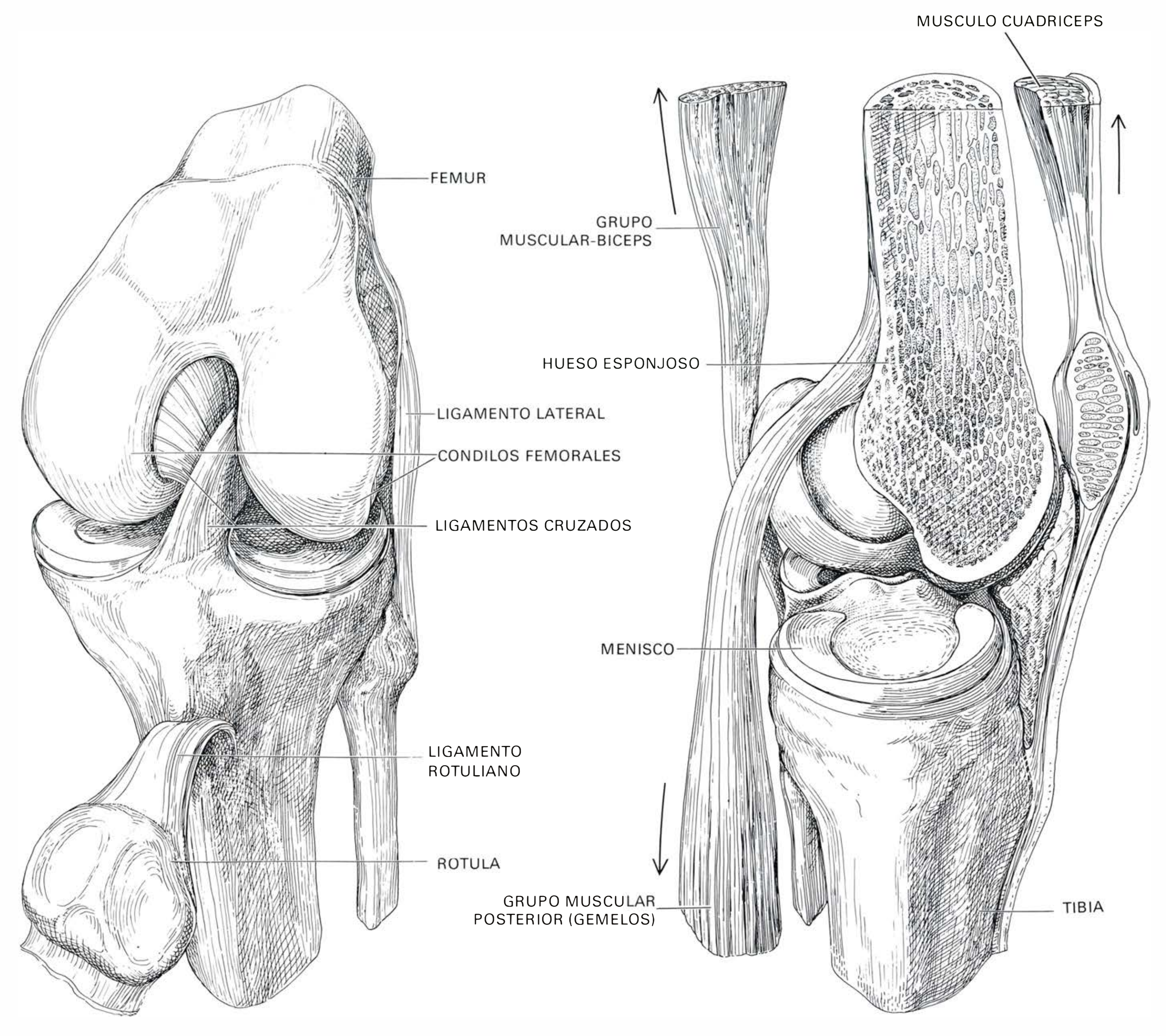
Las superficies de soporte de la rótula, la tibia y el fémur están revestidas por una capa relativamente delgada de cartílago, que es un tejido elástico. El cartílago es lubricado por el líquido sinovial, confinado en la cavidad articular gracias a una cápsula. Aunque quedan muchos interrogantes acerca de los mecanismos fisiológicos de lubricación articular, puede afirmarse que la acción conjunta del líquido sinovial y del cartílago proporciona a la articulación un mecanismo suave, duradero y de escaso rozamiento que ni las mejores articulaciones artificiales pueden igualar.

El uso de injertos metálicos para reparar o fortalecer el esqueleto humano es quizá tan antiguo como la misma cirugía. No obstante, el cuerpo plantea serias exigencias al metal. Hubo que esperar a nuestro siglo para que se definieran los requisitos de biocompatibilidad, esterilidad, resistencia mecánica y resistencia a la corrosión. Sólo un pequeño número de metales han demostrado ser satisfactorios. En los Estados Unidos, la mayoría de las prótesis se hacen de acero inoxidable o de una aleación de cobalto, cromo y molibdeno. El uso del titanio es cada día más frecuente.

A principios de la década de 1950 las investigaciones sobre tolerancia de materiales habían avanzado lo suficiente para que B. Walldius, del Instituto Carolino de Estocolmo, pudiera conseguir por primera vez la aplicación, con éxito, de una prótesis de rodilla. Esta prótesis, de tipo gozne, se construyó primero de material acrílico y más tarde de una aleación de cobalto, cromo y molibdeno. El dispositivo se fija al hueso mediante un cuidadoso modelado y ajuste de los componentes a los extremos de la tibia y del fémur y mediante tallos que se introducen en la cavidad medular de ambos huesos. La prótesis de Walldius ha aliviado el dolor y ha devuelto la función a muchos pacientes artríticos. Las causas más frecuentes de fracaso de esta prótesis han sido, o bien que se ha soltado del hueso, o bien que se han inflamado los tejidos debido al desgaste (metal contra metal) de las superficies de soporte. Con modificaciones que reflejan los progresos técnicos en cuanto a materiales para superficies articulares y a fijación de la prótesis, el diseño básico, de gozne fijo, constituye todavía el método de elección para algunos cirujanos como sustitución de rodillas gravemente afectadas.

El trabajo de Charnley sobre prótesis de cadera ha originado dos importantes

ESTRUCTURAS ESQUELETICAS implicadas en el funcionamiento de la pierna. La cadera y rodilla izquierdas se muestran provistas de prótesis. La creación de las articulaciones artificiales de rodilla se ha debido en parte al satisfactorio avance logrado en articulaciones artificiales de cadera.



RODILLA IZQUIERDA NORMAL, vista de frente y de perfil. En la visión frontal, la rodilla está ligeramente flexionada, mostrando el interior de la articulación y la disposición de los ligamentos. La visión lateral ilustra los principales grupos musculares que controlan la rodilla. Los músculos posteriores producen la flexión y el músculo cuadriceps la extensión. En la visión lateral se ha recortado parte del fémur.

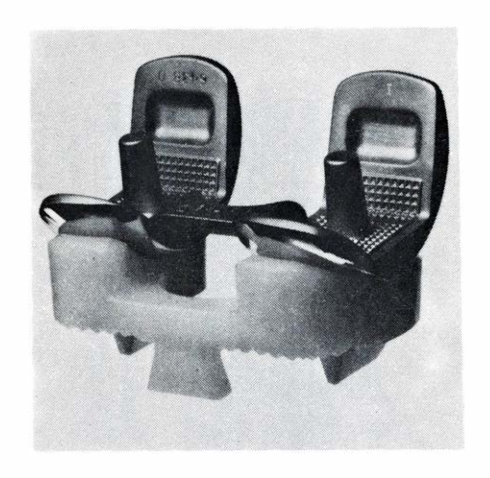
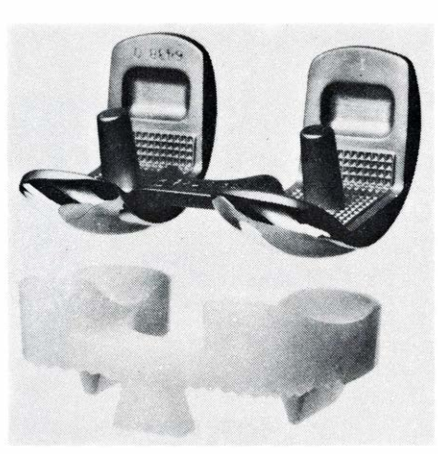
mejoras técnicas. Una consiste en la creación de superficies soportantes en las que se usa conjuntamente metal pulido y polietileno de elevado peso molecular. Con esta combinación se logran superficies articulares de bajo coeficiente de rozamiento de gran resistencia al desgaste, que dejan pocos residuos (que son, además, biológicamente inertes).

La segunda aportación de Charnley fue el uso de un cemento, el polimetil metacrilato, para fijar al hueso los componentes de una prótesis. El polimetil metacrilato no es un simple pegamento o adhesivo que actúe químicamente, sino un verdadero cemento de acción mecánica. Se utiliza en forma de pasta, por lo que puede fluir en torno a salientes y en el interior de cavidades óseas. Cuando fragua, forma una sujeción trabada mecánicamente.

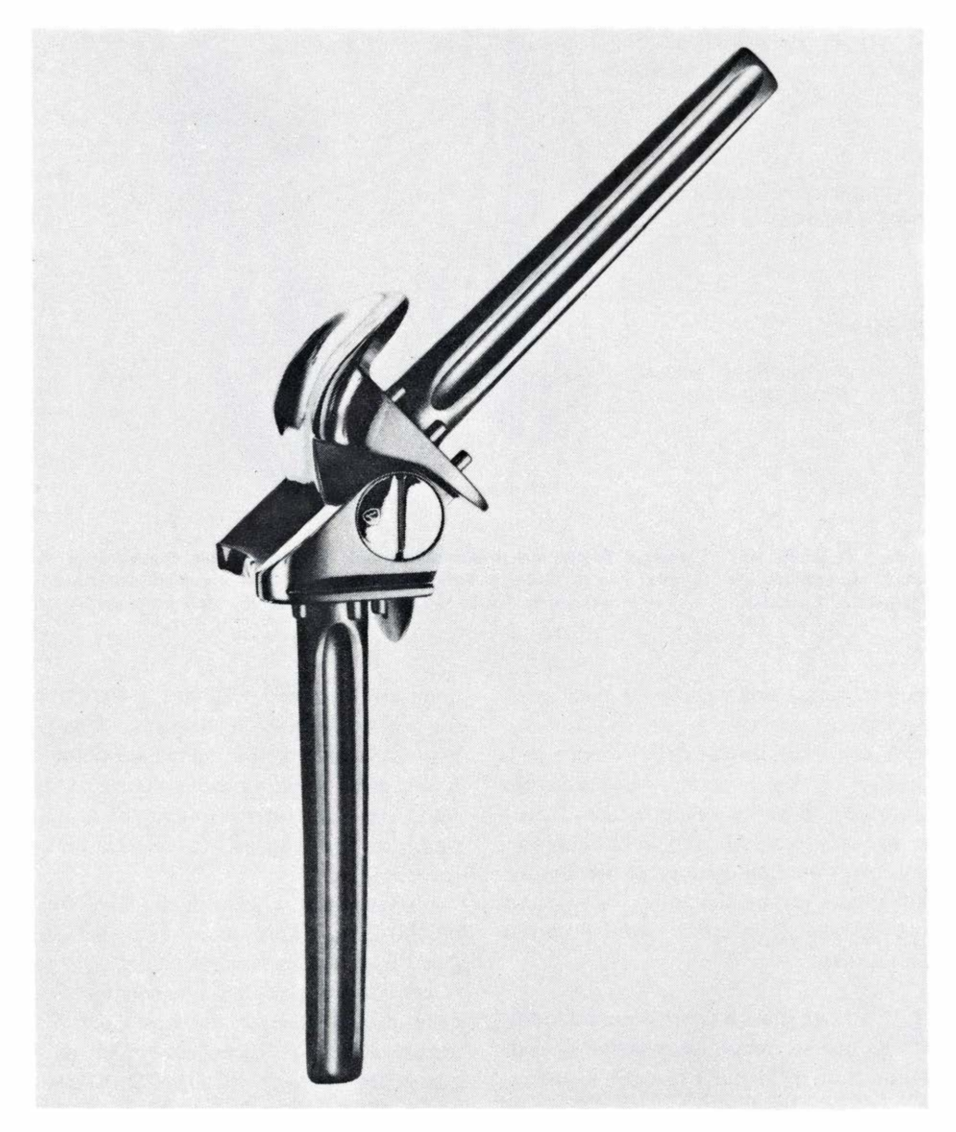
A estas mejoras se debe el éxito de la primera cadera total, realizada por Charnley. Sus dos componentes fueron la meseta tibial cóncava de polietileno y un correspondiente componente femoral metálico. Ambos elementos se fijaron a sus huesos respectivos con polimetil metacrilato.

El primer diseño de prótesis de rodilla en que se usaron estos adelantos se introdujo en 1971 por Frank H. Gunston, cirujano ortopédico asociado al Hospital General de Winnipeg y a la Universidad de Manitoba. La rodilla de Gunston es un dispositivo de carril y patín (deslizante). Los patines, hechos de metal, se unen con cemento al fémur, y los carriles, de polietileno, se cementan a la tibia. Los carriles son rectos de delante hacia atrás y ligeramente cóncavos de uno a otro lado; el movimiento de los patines en los carriles es relativamente irrestricto.

En los últimos años se ha trabajado mucho para definir la rodilla artificial ideal. Entre los temas debatibles figuraron la cantidad de movimiento relativo que debe proporcionarse entre los componentes femoral y tibial, los materiales que deben usarse en las superficies de soporte, la integridad estructural de los materiales plásticos, la repavimentación de la articulación rotulofemoral, la cantidad de hueso que debe quitarse y los criterios de selección de pacientes.

LA RODILLA GEOMETRICA es representativa de una clase de rodillas artificiales no articuladas, es decir, que carecen de articulación mecánica entre los componentes femoral y tibial. La prótesis se muestra montada *(izquierda)* y desmontada *(derecha)*. La estabilidad de este dispositivo depende de los ligamentos y músculos del paciente. De hecho, la rodilla no articulada proporciona al paciente nuevas superficies de apoyo en esa articulación. El componente femoral de la rodilla geométrica posee dos superficies de apoyo hechas de metal; las correspondientes piezas cóncavas del componente tibial son de polietileno. El dispositivo permite una flexión de 130 grados.



PROTESIS DE WALLDIUS, una de las primeras articulaciones artificiales de rodilla. La introdujo, a principios de la década 1950-59, Börje Walldius, del Instituto Carolino de Estocolmo. Por ser del tipo de gozne fijo, la prótesis de Walldius permite una amplitud de movimiento algo menor que la de una rodilla natural. Se fabrica ya con una aleación de cobalto, cromo y molibdeno.

Se calcula que, en todo el mundo, se dispone de más de 80 modelos de prótesis de rodilla. La mayoría difieren poco en su configuración y su función, por lo que conviene dudar acerca de esta proliferación. No obstante, el hecho pone de manifiesto los intensos esfuerzos dedicados al diseño de las prótesis de rodilla.

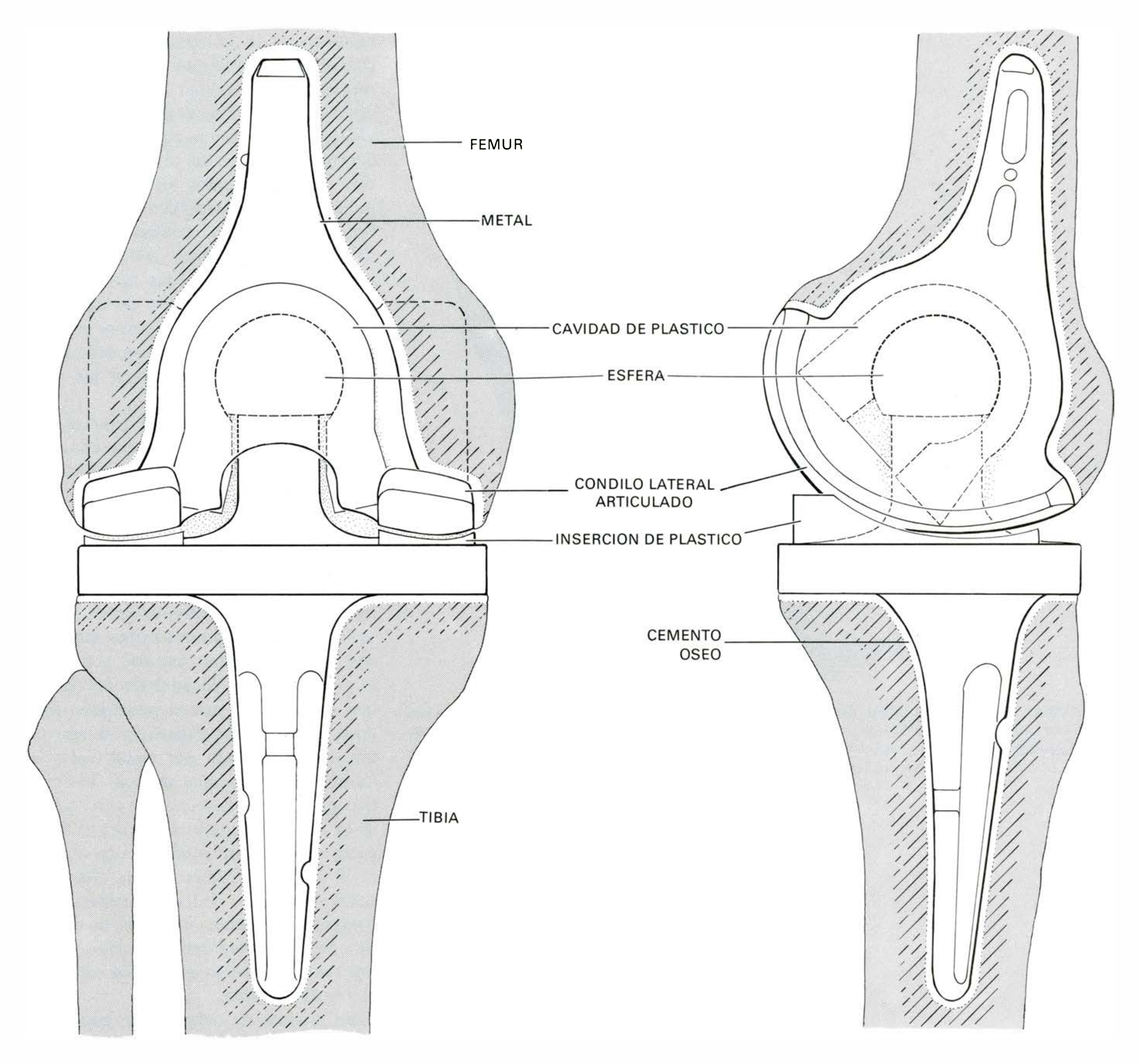
Centraremos nuestra atención sobre dos modalidades generales de sustitución de la rodilla. La primera, que incluye el grupo más numeroso, es el dispositivo no articulado o no vinculado. Su principal finalidad es la sustitución de las superficies articulares lesionadas o patológicas de la rodilla. Tales dispositivos no poseen goznes, enlaces mecánicos o similares entre los componentes femoral y tibial. En su lugar, la estabilidad de la articulación se consigue a expensas de los ligamentos y músculos del paciente con la ayuda, muchas veces, de la conformación geométrica de la prótesis. A esta modalidad responde la rodilla de Gunston.

El segundo tipo es el dispositivo articulado. Proporciona un vínculo mecánico restrictivo entre los componentes femoral y tibial. La rodilla de Walldius es un ejemplo clásico.

Uno de los dispositivos no articulados más empleado en los Estados Unidos es la rodilla geométrica, producto de trabajos de investigación ortopédica realizados en diversas instituciones. Aunque conceptualmente es similar al modelo de Gunston, consta de dos elementos en lugar de cuatro. Los problemas quirúrgicos de implantación y alineación son menores. Las superficies articulares son más congruentes, y proporcionan áreas más amplias de soporte y cierta restricción del movimiento.

El elemento femoral consta de dos superficies metálicas soportantes sujetas a los cóndilos, hechas de una aleación de cobalto, cromo y molibdeno, y tienen forma de esfera truncada. El componente tibial, de polietileno, posee superficies esféricas correspondientes de un diámetro ligeramente superior a su oponente femoral. Esta prótesis permite hasta 130 grados de flexión (la rodilla normal consigue unos 135 grados), así como pequeños movimientos de rotación lateral y axial.

Ambos componentes poseen un puente que conecta, por delante, sus mitades simétricas. Por tanto, cada uno de los componentes, visto desde arriba, tiene forma de *U*. Ello permite al cirujano conservar los ligamentos laterales y cruzados. (La anatomía de la rodilla permite implantar cualquier tipo de prótesis sin quitar los músculos.)



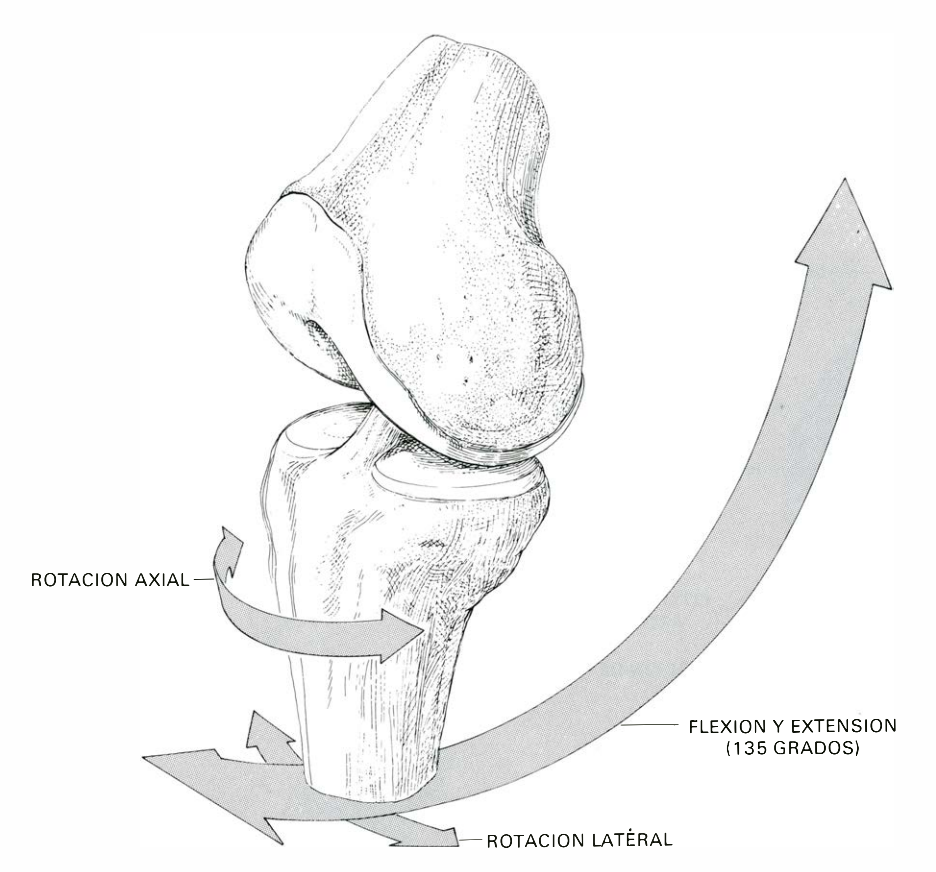
SISTEMA DE FUNCIONAMIENTO de una articulación artificial esferocéntrica de rodilla, indicado en dos cortes, frontal y lateral. Se trata de una prótesis articulada por un dispositivo de bola y concavidad (mecanismo convexo-cóncavo). Los cojinetes son de metal sobre plástico.

Para colocar una rodilla geométrica el cirujano comienza resecando una fina capa de las superficies de apoyo de la rodilla enferma. Los extremos óseos deben ser remodelados para que se adapten correctamente a los componentes de la prótesis, que se fijarán mediante polimetil metacrilato. Dichos componentes van provistos de tallos, estrías y colas de milano para facilitar la acción del cemento.

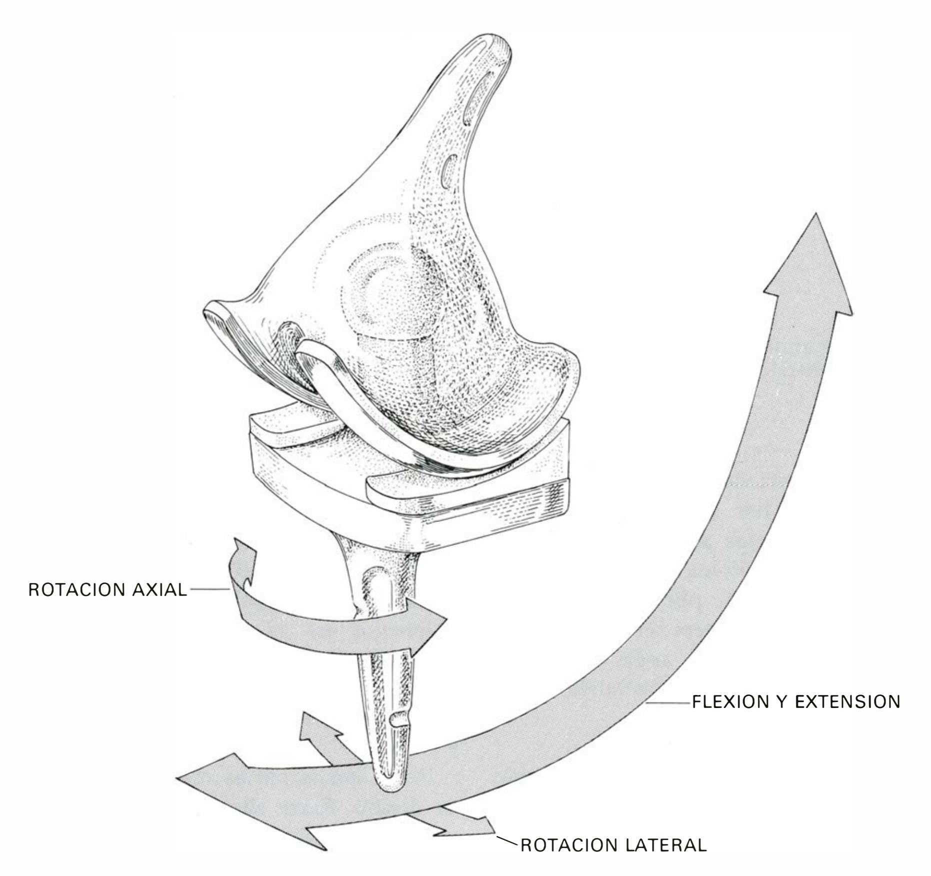
La rodilla geométrica proporciona al paciente unas superficies de apoyo repavimentadas en su articulación artrítica. Su empleo durante casi 10 años ha puesto en evidencia que la rodilla geométrica puede aliviar el dolor y permitir a su portador una actividad física aceptable y, por supuesto, muy superior a la que tenía antes de la intervención. Los dispositivos se averían en pequeña proporción, por lo común por soltarse o deformarse el componente tibial. En tales casos, suele ser posible la implantación de otro modelo, como la rodilla esferocéntrica.

La rodilla esferocéntrica es el resultado de un trabajo iniciado por nuestro equipo en 1971 para diseñar una rodilla artificial articulada. Los modelos no articulados, como la rodilla geométrica, resultan útiles para pacientes cuyo principal problema radica en el desgaste de las superficies articulares y cuya rodilla enferma conserva buenos ligamentos. Para las rodillas artríticas muy incapacitadas, con gran deformidad e inestabilidad, el único modelo que se usaba ampliamente era el dispositivo de gozne fijo, del tipo Walldius. En vista de los progresos técnicos conseguidos, pensamos que una alternativa era correcta y posible.

Nos fijamos varias metas y criterios de diseño. Entre ellos, habíamos de diseñar un enlace mecánico articulado entre los componentes femoral y tibial que proporcionara estabilidad intrínseca, con independencia de los ligamentos; que permitiera la movilidad triaxial ca-



MOVIMIENTOS TRIAXIALES de que es capaz una rodilla normal, algo esquematizados. Además de flexionarse hasta 135 grados, la rodilla puede realizar una rotación axial, como cuando se hace oscilar el pie a uno y otro lado a la altura del tobillo, y una rotación lateral, lo cual significa que la tibia gira respecto del fémur. Esta rotación se produce al flexionar y estirar la pierna.



AMPLITUD DE MOVIMIENTOS que permite la rodilla esferocéntrica. En el caso ideal, el paciente provisto de dicha prótesis puede lograr 120 grados de flexión y extensión (alrededor del eje *Y*), 30 grados de rotación axial (eje *Z*) y cinco de rotación lateral (eje *X*). El movimiento lo controlan principalmente los carriles y patines, que poseen un conjunto de radios de curvatura variables.

racterística de la rodilla normal; que controlara la amplitud de movimientos mediante deceleración gradual, evitando bruscas detenciones sucesivas del movimiento; que pusiera en contacto metal contra polietileno en todas las superficies de apoyo; todos los elementos de polietileno irían reforzados con metal a fin de minimizar la tendencia del plástico a la deformación; usaríamos polimetil metacrilato para las fijaciones; se garantizaría la sustitución de cualquier superficie de polietileno que se gastase y debíamos conseguir un dispositivo que el cirujano pudiese implantar sin dificultades indebidas.

El criterio de movilidad triaxial significaba que la unión convexo-cóncava sería la mejor manera de lograr un enlace mecánico entre los componentes. Un modelo de carril y patín, que tuviese la misma relación con el modelo convexo-cóncavo que mantiene un conjunto de balancines con una embarcación, proporcionarían la amplitud y los límites de su movilidad. Conseguimos ambos aspectos con una concavidad femoral y dos inserciones tibiales ideadas de tal forma que dondequiera que el polietileno formase parte de una superficie de apoyo, estuviese soportado por metal *(véase la ilustración de la página opuesta).* El objetivo de lograr la deceleración controlada de la extensión completa de la rodilla se cumplió finalmente mediante una correcta situación del mecanismo cóncavo-convexo respecto de los balancines, cuyo conjunto de cambios de radios de curvatura fue cuidadosamente escogido y que, por tanto, permitían una acción semejante a la de una leva.

En pruebas de carga sobre huesos de cadáver descubrimos que tanto las prótesis del tipo Walldius como la esferocéntrica podían soportar, antes de romperse, cargas superiores a las de una rodilla normal. De especial interés eran las características de carga dinámica de cada dispositivo al detener el movimiento al final de la extensión. La carga puede medirse como energía acumulada.

En nuestros laboratorios hemos estudiado diversos diseños del modelo esferocéntrico, a fin de emparejar estrechamente el carácter de la acumulación de energía de la prótesis con dicho carácter en una rodilla normal. Nuestro modelo final de la prótesis de rodilla proporciona cargas de acumulación de energía de alrededor de un tercio de las comprobadas con el modelo de Walldius, cuyos topes son de metal contra metal. Por cuanto las cargas que participan en la detención de la extensión se transmiten

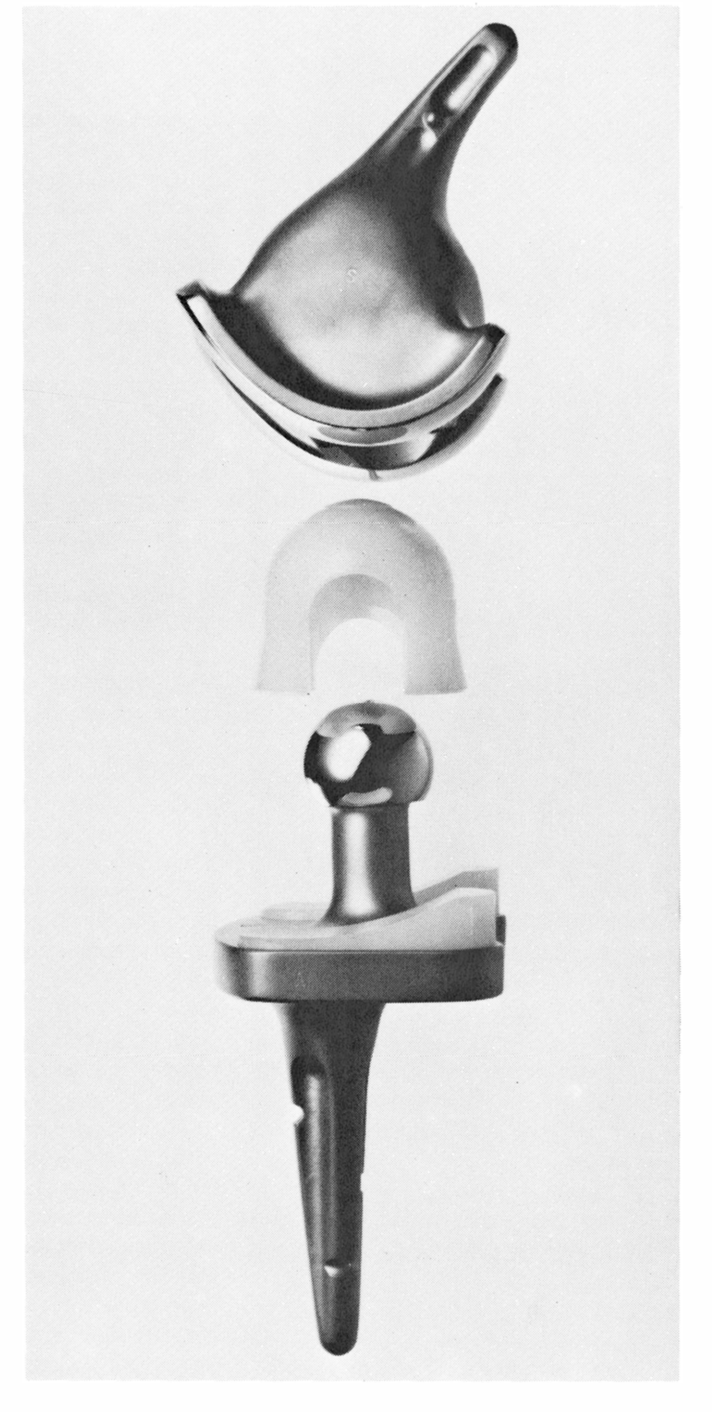
a las superficies a las que la prótesis va fijada con cemento, parece probable que el modelo esferocéntrico tendrá muchos menos problemas de aflojamiento que la prótesis de Walldius y los demás dispositivos de gozne fijo, restringidos de manera análoga.

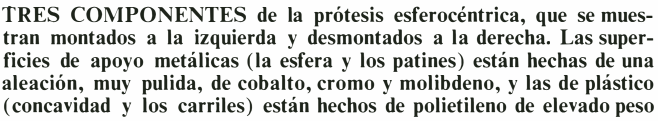
En el año 1973 estábamos preparados para iniciar las pruebas clínicas de la rodilla esferocéntrica. El criterio para la selección de pacientes fue (y continúa siendo) el siguiente: deformidad importante y gran inestabilidad o fracaso de una prótesis anterior. Nuestro primer ensayo abarcó 22 pacientes y 25 prótesis. El control postoperatorio de estos pacientes nos permitió comprobar que experimentaron un alivio del dolor y obtuvieron una capacidad de actividad física mucho mayor.

Hasta ahora se han implantado más de 120 prótesis esferocéntricas en el Centro Médico de la Universidad de Michigan. Sin embargo, esta técnica debe seguir considerándose experimental. Los éxitos en cuanto a movilidad triaxial y deceleración controlada parecen haber disminuido el problema de desinserción de los componentes, pero solamente una ulterior experiencia con este dispositivo demostrará si la desinserción puede evitarse a largo plazo.

Una encuesta reciente sobre tecnolo-







TRES COMPONENTES de la prótesis esferocéntrica, que se muestran montados a la izquierda y desmontados a la derecha. Las superficies de apoyo metálicas (la esfera y los patines) están hechas de una aleación, muy pulida, de cobalto, cromo y molibdeno, y las de plástico (concavidad y los carriles) están hechos de polietileno de elevado peso molecular. Para implantar la prótesis, el cirujano coloca el componente femoral y luego el tibial. La pieza cóncava de polietileno se pone sobre la esfera y se presiona encajándola en el componente femoral en donde unos cerrojos la fijan en su lugar. En la prótesis esferocéntrica, todas las superficies de apoyo hechas de plástico están soportadas por metal.

gía de rehabilitación, realizada en la Northwestern University, indicó que en el año 1976 se habían llevado a cabo 60.000 reemplazos de cadera y 30.000 de rodilla en Estados Unidos. Los cirujanos, al rellenar un cuestionario que formaba parte de la encuesta, indicaron que si se contara con prótesis más de fiar, efectuarían 1,12 veces más reemplazos de cadera y 1,76 veces más de rodilla. Estas cifras sugieren que se precisará un importante esfuerzo para que la sustitución de la rodilla humana pueda considerarse como algo más que un éxito con algunas reservas.

A nuestro parecer ese esfuerzo debe realizarse en dos aspectos. El primero se relaciona con el diseño de rodillas artificiales. Están apareciendo nuevos y prometedores materiales protéticos, tales como metales, polímeros y cerámicos. Con ellos se mejoraría el rendimiento mecánico, y se tendría más libertad al diseñar los dispositivos. Las investigaciones en curso sobre la carga de rodillas normales y patológicas llevará a mejores pruebas de los nuevos diseños en el laboratorio.

El segundo aspecto que se precisa perfeccionar es la fijación del dispositivo al hueso. La desinserción constituye la causa más frecuente de fracaso de las prótesis de rodilla (aproximadamente un 10 por ciento en los dos primeros años). En el momento actual no se sabe lo suficiente sobre este fenómeno como para predecir cuáles son los pacientes que más probablemente lo experimentarán. Es evidente que la frecuencia y gravedad de la desinserción aumentan con el tiempo, y puede decirse que es muy probable que la complicación se produzca en pacientes que, por su actividad física o su peso, exigen mucho a sus prótesis.

Hemos estado estudiando este problema en el laboratorio mediante la aplicación de diferentes cargas a uniones de cemento y hueso. Los resultados han demostrado que el polimetil metacrilato se fija mejor en el hueso esponjoso (material esponjoso de los extremos de los huesos largos) que en el hueso cortical (el material más duro existente en los extremos). Hemos comprobado también amplias variaciones en el comportamiento del tejido esponjoso de diferentes individuos, de distintos huesos de un mismo sujeto e, incluso, de distintas zonas de un mismo hueso.

Estos estudios se han realizado con huesos de cadáver, mientras que el hueso vivo constituye un tejido activo, capaz de desarrollar resistencia y remodelarse en respuesta a las cargas. Junto con otros investigadores creemos que el aflojamiento de las prótesis se relaciona con ese mecanismo. Una vez implantada la prótesis, el hueso adyacente al cemento se reabsorbe y se remodela activamente para adaptarse a las nuevas cargas. Se presenta entonces un frágil equilibrio en el que a veces domina la reabsorción y se produce la desinserción de la prótesis. Solamente en experimentos bien planificados sobre animales vivos se puede investigar con mayor fruto el problema de las interacciones entre hueso y cemento.

Otros investigadores están estudiando dos métodos de fijación que aparecen como alternativos. Uno de ellos está basado en las propiedades piezoeléctricas del hueso. Los potenciales eléctricos generados por las presiones que fisiológicamente soporta el hueso o por otras que puedan aplicársele artificialmente, quizás estimulen el crecimiento óseo y, en consecuencia, refuercen la fijación de una prótesis. El segundo enfoque consiste en implantar prótesis con superficies porosas que, al facilitar el crecimiento en el interior del tejido óseo neoformado, permiten la fijación directa de la prótesis. Matrices de cerámica, polímeros y metales, de dimensiones adecuadas, han mostrado amplias posibilidades al respecto.

La sustitución de la articulación de la rodilla humana constituye un campo de trabajo en continua y rápida evolución. Las 30.000 prótesis implantadas en 1976, especialmente en pacientes ancianos y con invalidez artrítica crónica, han aliviado el dolor y han proporcionado más capacidad funcional a la mayoría de sus portadores. Puede contarse con que varias cosas harán que estas técnicas sigan progresando. Entre ellas: el perfeccionamiento de los materiales constitutivos de las prótesis, la identificación de técnicas de laboratorio adecuadas para comprobar los diseños y el refinamiento de las pruebas de deambulación para evaluar al paciente antes y después de la intervención. Sin embargo, es preciso resolver el problema de la desinserción de los componentes de la prótesis, principal causa del fracaso clínico de la intervención, antes de que la técnica de sustitución de la articulación de la rodilla se generalice más ampliamente. Y aún entonces, los pacientes, los cirujanos y los técnicos mecánicos no deberían olvidar que la rodilla humana es un notable producto de la ingeniería evolutiva. Aunque se la puede imitar con éxito creciente, jamás podrá ser igualada.

Ciencia y sociedad

El citoesqueleto

A la célula viva solía considerársela en buena medida como una bolsita diminuta rellena de un líquido en el que interaccionan libremente enzimas y otras substancias. Con los años, se advirtió que el interior de la célula estaba organizado, con rigor mucho mayor, en órganos y compartimientos membranosos. Pero va haciéndose cada día más evidente que las células poseen asimismo un intrincado esqueleto interno compuesto de redes de proteína fibrosa. Aunque con anterioridad se había caracterizado a diversos tipos de fibras subcelulares mediante el microscopio electrónico, su configuración general se reveló por primera vez con claridad cuando se usaron anticuerpos específicos para rotular las fibras con marcadores fluorescentes, de modo que pudiera identificárseles con el microscopio óptico. Dichos experimentos mostraron que los tres tipos principales de fibras subcelulares –los microtúbulos, los microfilamentos y las fibras intermedias– están organizadas en redes que se extienden por todo el interior de la célula.

Los microtúbulos, de unos 250 angströms de diámetro, se extienden radialmente hacia fuera desde el núcleo celular siguiendo vías rectas o curvas hasta cerca de la superficie de la célula. Ayudan a mantener la forma de ésta y guían asimismo el movimiento de orgánulos tales como las mitocondrias y lisosomas de una a otra parte de la célula. Al inicio de la división celular, el sistema microtubular está organizado formando el haz mitótico, que participa en la separación de los cromosomas dobles conforme se divide la célula. Los microfilamentos, de unos 60 angströms de diámetro, suelen estar relacionados con el interior de la membrana celular. Se componen fundamentalmente de actina, proteína contráctil, y también de otras proteínas que se encuentran asimismo en las fibras musculares. Los microfilamentos están organizados en mallas o haces, ambos importantes para la motilidad celular. Los filamentos intermedios miden unos 100 angströms de diámetro y forman una red ondulada y entretejida dentro de la célula, distinta de otros sistemas fibrosos. Los tipos de fibra son estructuras dinámicas que están constantemente armándose, desarmándose y reordenándose en las células de motilidad activa.

Al objeto de examinar la estructura de la matriz citoplasmática con mayor detalle, Keith R. Porter y sus colegas de la Universidad de Colorado, en Boulder, usaron uno de los dos microscopios electrónicos de alta tensión actualmente existentes en los Estados Unidos. Este instrumento acelera electrones en un campo con una diferencia de potencial de un millón de voltios, dotándolos de energía cinética suficiente como para penetrar en células intactas y muy aplanadas. La imagen resultante es análoga a una radiografía, y revela la estructura interna de la célula sin necesidad de rebanar la célula en finos cortes, como es preciso hacerlo para examinarla con los microscopios electrónicos de menor tensión.

Cuando Porter examinó células cultivadas de pulmón de feto humano (que son extraordinariamente planas) con el microscopio electrónico de alta tensión, halló que la totalidad de su citoplasma estaba lleno de una red de fibras cual encaje, que no había sido vista hasta entonces. Las fibras eran más delgadas que los microfilamentos –de 30 a 60 angströms de diámetro–. Debido a que la red fibrilar recordó a Porter la estructura trabecular del hueso esponjoso, la denominó "sistema microtrabecular". El tramado parece servir como una especie de andamio intracelular al que van anclados los diversos orgánulos, como también los tres sistemas principales de fibras. Porter cree que funciona como una especie de musculatura celular, que redistribuye y reorienta constantemente los orgánulos, conforme la célula desempeña sus actividades, y que coordina las funciones de los tres sistemas principales de fibras. El movimiento de los componentes celulares a través del tramado se logra, al parecer, mediante la contracción y extensión de cada filamento microtrabecular, almacenándose energía cuando están extendidos.

La hulla como alternativa

Se ha calculado que el carbón representa hasta un 90 por ciento del total de reservas de combustible fósil de los Estados Unidos, y no obstante ahora satisface solamente alrededor del 18 por ciento de las necesidades energéticas de ese país. En vista de los menguantes abastecimientos nacionales de petróleo y gas natural, la creciente dependencia respecto de los combustibles fósiles importados y la improbabilidad de que otras fuentes de energía puedan contribuir significativamente al presupuesto energético de los Estados Unidos a lo largo de los próximos 25 años, la atención se ha centrado en la hulla como un componente principal de su estrategia energética hacia el futuro. De acuerdo con esta perspectiva, la Administración Carter ha incluido en su plan energético nacional, ahora ante el Congreso, la propuesta de duplicar la producción hullera en los próximos diez años: de 665 millones de toneladas en 1976 a 1200 millones de toneladas hacia 1985. ¿Qué probabilidades hay de que Estados Unidos pueda lograr ese aumento de su producción hullera anual?

Según un voluminoso informe preparado por la Oficina General de Cuentas (GAO) para el Congreso norteamericano, que lleva por título *Desarrollo de la hulla en Estados Unidos. Promesas e incertidumbres,* las perspectivas no son alentadoras. Su conclusión es la siguiente: "de hecho será muy difícil alcanzar los 1000 millones de toneladas para 1985." Los detallados análisis que respaldan esta conclusión están expuestos en el informe, que consigna "el estado actual, perspectivas y principales asuntos a debatir en el desarrollo hullero de los Estados Unidos desde el punto de vista de la demanda, el abastecimiento, la producción, el transporte, repercusiones ambientales y socioeconómicas y la posición del país en el mercado hullero mundial."

El descenso del carbón desde su posición preeminente de hace 50 años, cuando proporcionaba el 80 por ciento de la energía consumida en los Estados Unidos, se ha frenado, según parece, por la reciente renovación del interés en sus posibilidades. Sin embargo, según lo deja en claro el informe antedicho, "el renovado interés por el carbón como fuente de energía es asunto más de necesidad que de elección. Si fuese un asunto estrictamente opcional, continuaría la decadencia del carbón respecto de otros combustibles. No se le preferiría al petróleo y al gas por varios motivos fundamentales: se lo excava, no se lo bombea y por tanto es más difícil y peligroso extraerlo de la tierra; es más voluminoso y por tanto más difícil de transportar y manipular; más sucio y por tanto oca-

siona más contaminación cuando se lo quema". De hecho, lo mejor que puede decirse acerca del carbón es, al parecer, "que hay muchísimo", especialmente respecto de los recursos nacionales de gas y petróleo.

Sin embargo, los precios recientes del carbón no han sido tan atractivos como los de otros recursos energéticos. Entre otros motivos expuestos en el citado informe figuran: (1) las incertidumbres acerca de las normas ecológicas, (2) los posibles aumentos en los costos de capital y de explotación ocasionados por los requisitos de control ambiental, (3) los problemas de transporte y almacenamiento, y (4) las ventajas relativas de la energía nuclear en cuanto a costos.

El plan energético de la Administración Carter aborda algunas de las limitaciones para un mayor uso del carbón, pero en opinión de los analistas de la GAO omite otras, principalmente las referentes a transporte y productividad. Al llegar a su propia versión de los límites físicos y económicos de "la solución hullera", la GAO se inclina por la conclusión de que "para duplicar la producción de hulla hacia 1985 habrían de actuar tantos factores interconexos que la GAO no cree que pudiese suceder tal cosa; para mencionar tan sólo dos, los fabricantes de maquinaria tendrían que satisfacer oportunamente los pedidos y las compañías mineras habrían de tener la previsión y el capital suficientes como para abrir nuevas minas cuando se precisara la producción adicional".

Durante el período entre hoy y 1985, continúa el informe, "el carbón no sólo es inelástico en su oferta, sino también en su demanda, en el sentido de que las compañías eléctricas y usuarios industriales no van a comprar carbón si no pueden consumirlo. Hay largos períodos de preparación, tan sólo en construir e instalar calderas en centrales y fábricas existentes, para no hablar del tiempo que ha de pasar para proyectar y construir centrales energéticas a carbón de nueva planta".

En un futuro a plazo medio (de 1985 a 2000), se añade en el informe, la hulla continuará estando limitada por la demanda. "Las posibilidades de una sustitución directa del petróleo y del gas son muy limitadas en cuanto al total de la economía. Las de sustitución indirecta mediante electricidad producida quemando carbón, si bien son más prometedoras, también están limitadas por factores económicos y por el estado actual de la tecnología industrial y del transporte".

A más largo plazo (más allá del año 2000), a los analistas de la GAO les parece que el carbón volverá a estar constreñido tanto por la oferta (especialmente tratándose de hulla con poco azufre y para metalurgia) como por la demanda; y opinan que las perspectivas a muy largo plazo para una mayor demanda de carbón "se apoyan en la esperanza en que los gases y líquidos derivados del carbón lleguen a ser económicos y ecológicamente seguros".

Según el informe de la GAO, no hay duda de que "la hulla satisfará gran parte del futuro energético de los Estados Unidos". Pero los "cambalaches" exigidos serán costosos, "especialmente en vidas humanas y enfermedades". Si en último análisis se decide que el costo de usar más allá de cierta proporción de carbón es excesivo y que una mayor importación no es una alternativa soportable, entonces, se afirma en el informe, a los Estados Unidos le quedan sólo dos opciones entre el presente y el año 2000: (1) acelerar el aumento de la energía nuclear convencional, de modo que la electricidad así generada sustituya en lo posible al petróleo y al gas, o (2) perfeccionar la conservación de la energía mediante una mayor eficiencia y un menor consumo.

Una revolución protocientífica

¿Cuándo empezaron los filósofos a concebir el universo como algo natural que podía comprenderse mediante una investigación racional y sistemática, y no como entidad sobrenatural a la cual contemplar sólo con pasivo pasmo? Aunque la aparición del pensamiento racional suele asignarse a la revolución copernicana y al Renacimiento, en un artículo reciente de la revista historiográfica *Isis*, Tina Stiefel expone documentos de principios del siglo XII que prueban que "cuatro siglos antes del inicio de la ciencia moderna, un pequeño grupo de europeos occidentales [...] idearon un concepto de la ciencia como disciplina aparte, y avanzaron bastante hacia la formulación de una metodología digna de crédito para practicarla."

La prueba principal viene aportada por escritos de varios cosmólogos de esa época, en especial Adelardo de Bath, Thierry de Chartres y Guillermo de Conches. Todos ellos eran estudiosos de Platón, de las obras de autores árabes y de la lógica aristotélica (que hasta 1150 era la única parte de las obras científicas de Aristóteles conocida en Europa). "Tal vez no nos hemos percatado con bastante claridad de la medida en que la práctica de la lógica aristotélica transformó la creencia de los hombres en su propia facultad racional", escribe Stiefel. "Profundamente confiados en su propia preparación en la nueva dialéctica, los cosmólogos extendieron su confianza al poder de la razón humana para penetrar en los secretos de la naturaleza."

Los cosmólogos del siglo XII expresaron firmemente su convicción de que el universo era básicamente racional. En su libro *Quaestiones naturales,* Adelardo de Bath escribió: "Debemos suponer que toda la naturaleza se asienta sobre un fundamento seguro y lógico"; todo lo que existe tiene un motivo lógico para existir, y el hombre, parte de la creación, está dotado para entender esos motivos. Pero esa comprensión puede provenir solamente del cultivo y adiestramiento de la mente de un modo lógico. Guillermo de Conches intentó definir la ciencia como la parte del conocimiento que se funda en cosas corpóreas, en oposición a la metafísica y a la teología, que se ocupaban de cuestiones sobre la esencia y la substancia invisibles. Stiefel observa: "Definir la ciencia era menos fácil de lo que podríamos suponer, y frente a la resistencia, comúnmente sentida, a fijar límites al ámbito de la teología, se requería cierto valor para aventurar tan siquiera una definición".

Esos cosmólogos del siglo XII llegaron también a la idea de explicación científica teórica; aunque no formularon completamente una metodología sistemática para llegar a tal explicación, sí reconocieron sus partes componentes. Adelardo de Bath y Thierry de Chartres propugnaban técnicas empíricas; la naturaleza debía estudiarse mediante un examen racional de las pruebas presentadas por los sentidos. Juan de Salisbury, discípulo de Guillermo de Conches, expuso en su *Metalogicon* la función del razonamiento inductivo en la formulación de hipótesis, y el papel que la probabilidad cumple en las descripciones de sucesos naturales. Adelardo y otros tenían confianza en el poder del razonamiento matemático para deducir conocimientos acerca de la naturaleza.

Los cosmólogos pensaban asimismo que era indispensable que el filósofo natural abordara el estudio científico con una mentalidad específica, a saber, que debía empeñarse en lograr un estado de desapegada objetividad, a fin de liberar su capacidad útil de construir hipótesis de la tiranía de suposiciones aceptadas sin más y de las opiniones aceptadas acerca de la naturaleza. En opinión de Stiefel, "la noción de convertir la duda metódica en hábito es quizá la más sofisticada de las ideas [...] porque implica una peligrosa libertad del intelecto".

Según Stiefel, la revolución conceptual representada por la obra de Adelardo, Thierry y Guillermo ha pasado en gran parte inadvertida, porque en el siglo XII tuvo un influjo muy limitado. Su promulgación coincidió con el inicio del movimiento escolástico, que se ocupó de aplicar la lógica aristotélica a las creencias cristianas. Ensombrecida por el escolasticismo, la obra de los protocientíficos fue ignorada o tratada como herejía, y sus proponentes fueron ridiculizados si no perseguidos. Sin embargo, sus conceptos pesaron en eruditos del siglo siguiente, tales como Roberto Grosseteste y Roger Bacon, cuya influencia sobre la moderna revolución científica fue más directa.

Concluye Stiefel: "El que los cosmólogos avanzaran algo en el sentido de formular una nueva disciplina y delimitaran métodos practicables para abordar la ciencia, es una proeza que merece una mayor atención de los historiadores. Lo que en gran medida no se ha tenido en cuenta es el salto imaginativo que exigió esa obra, porque es fácil no advertir el gran coraje que supuso el dar ese salto: desde una sumisa aceptación de la naturaleza como necesariamente inescrutable para un espíritu inquisitivo, optimista y confiado a la vez... Por su valentía y originalidad, los cosmólogos son dignos de mención y sus empeños por construir la disciplina de la ciencia merecen atención y respeto."

HLA y enfermedad

La individualidad del tejido de seres humanos diferentes la afirman principalmente los productos proteicos de genes situados en cuatro *loci* en el brazo corto del cromosoma número 6: el complejo principal de histocompatibilidad, designado por *HLA*. Los genes *A*, *B* y *C* codifican los antígenos insertados en la membrana celular, y parece que el gen *D* participa en la fabricación de anticuerpos. Los cuatro genes son sumamente polimórficos, es decir, hay muchos alelos, o formas alternas, diferentes de cada uno. La diferencia resultante en los antígenos *A*, *B* y *C* es la que hace que el sistema inmune reconozca como tal al tejido extraño o ajeno y lo rechace. El polimorfismo del complejo, cuyos productos en individuos diferentes pueden distinguirse mediante pruebas de laboratorio, ha proporcionado a los investigadores de genética una importante herramienta para estudiar la relación entre determinados genes y determinadas enfermedades y la pauta según la cual se heredan las enfermedades que poseen un componente genético. En el *The New England Journal of Medicine* se ha publicado, junto con editoriales en que se valora el avance en este campo de investigaciones, unos recientes resultados de estudios en los cuales se relaciona al complejo *HLA* con la susceptibilidad a dos enfermedades.

En estudios de poblaciones se ha investigado más de 100 enfermedades en busca de pruebas de su asociación con antígenos *HLA*, específicos, y en más de la mitad de ellas se ha informado sobre resultados positivos. Tal como lo señalan en sus editoriales en la publicación antedicha, Leon E. Rosenberg y Kenneth K. Kidd, de la Facultad de Medicina de Yale, y James V. Neel, de la Facultad de Medicina de la Universidad de Michigan, las asociaciones pueden explicarse de tres modos. Determinados antígenos *HLA* podrían estar directamente implicados en la etiología de la enfermedad, la enfermedad podría ser común en un subgrupo étnico de la población estudiada en la cual determinados alelos *HLA* sean también comunes o la explicación podría ser un "desequilibrio de ligamiento": el gen que aumenta la susceptibilidad a la enfermedad podría estar ligado con desproporcionada frecuencia con alelos específicos en los *loci HLA*. Se estima que esta última explicación es la más probable para la mayoría de las asociaciones. Una asociación que sugieran los estudios de poblaciones puede confirmarse mediante estudios de las familias afectadas, y los estudios de familias pueden asimismo indicar el medio de efectuarse la herencia.

A la hemocromatosis, que es el desorden del metabolismo del hierro caracterizado por la excesiva deposición de hierro en los tejidos, se la ha relacionado con dos antígenos *HLA*, el *A*3 y *B*14. En otras palabras, los estudios de poblaciones han demostrado que las personas que lleven esos determinados antígenos (y los alelos que los codifican) tienen más probabilidades que otras de padecer hemocromatosis. Un informe de cuatro investigadores del hospital Ponchaillou, de Rennes, Francia, respalda ese descubrimiento y la herencia recesiva de la susceptibilidad a la hemocromatosis. Marcel Simon y sus colegas determinaron los haplotipos de personas con un mismo progenitor en familias en que dos o más de los hijos mostraban síntomas convincentes de la enfermedad. (Un individuo puede tener dos alelos diferentes en cada *locus*, en este caso los *loci HLA-A* y *HLA-B*; se hereda un alelo de cada progenitor. Los alelos heredados

de un progenitor constituyen el haplotipo y dos haplotipos en conjunto constituyen el genotipo. Ambos haplotipos se expresan en los genes *HLA:* los genes son "codominantes".)

Los investigadores descubrieron que una proporción significativa de los hijos afectados tenían ambos haplotipos en común, más frecuentemente el *A*3 y el *B*14, pero no siempre; muy pocos de los hijos no afectados de estas familias tenían un genotipo idéntico al de un hijo afectado. Tales resultados implican que la hemocromatosis declarada depende de la presencia de dos cromosomas homólogos específicos –que su modo de transmisión es recesivo (en cuanto es opuesto a la transmisión dominante, en la cual un gen correspondiente a un rasgo es suficiente para que tal rasgo se exprese). Dicha conclusión está respaldada por estudios de familias con un hijo afectado (en la mayoría de los casos los miembros no afectados tenían haplotipos diferentes) y por análisis estadístico de un grupo de personas no emparentadas afectadas de hemocromatosis. Los autores sugieren que la tipificación *HLA* en familias con hemocromatosis podría detectar la enfermedad antes de aparecer los síntomas y hacer posible el tratamiento preventivo.

Una enfermedad cuya base genética se ha debatido largamente es la diabetes mellitus. Parece ahora que la forma de aparición juvenil de la enfermedad (que parece ser clínica y genéticamente diferente de la forma de aparición adulta, menos grave) es ocasionada por una respuesta autoinmune a una infección vírica, y que la susceptibilidad a la respuesta está influida por uno o varios genes cercanos a los *loci HLA*.

En un artículo que acaba de aparecer recientemente en el *New England Journal,* Pablo Rubinstein, del New York Blood Center, y Nicole Suciu-Foca y John F. Nicholson del College of Physicians and Surgeons de la Universidad de Columbia, alegan, fundándose en estudios de familias, que un gen que predispone a la diabetes juvenil está específicamente relacionado con el *locus D* del complejo *HLA* y que la herencia es recesiva: que se precisan dos de tales genes para que la enfermedad se manifieste. Neel señala en su editorial que esa conclusión debía considerarse tentativa. Un motivo es la posible heterogeneidad de la enfermedad: la carencia de insulina que caracteriza a la diabetes puede estar ocasionada por defectos genéticos múltiples con distinto grado de penetratividad o expresión, y cuya incidencia varía en distintas poblaciones.

La estructura tridimensional del ARN de transferencia

Este ácido nucleico desempeña un papel fundamental en la traducción del código genético en la secuencia de aminoácidos de una proteína. La determinación de su estructura ha venido a clarificar la síntesis proteica

Alexander Rich y Sung Hou Kim

Es ya ampliamente conocido que las instrucciones para el ensamblaje y organización de un sistema vivo se hallan incorporadas en las moléculas de ADN contenidas en la célula viva. La secuencia de bases de nucleótidos a lo largo de la cadena de la molécula de ADN especifica la estructura de miles de proteínas, que son el material de construcción de la célula y los catalizadores de sus intrincadas reacciones bioquímicas. No obstante, una molécula de ADN es en sí algo parecido a una cinta magnetofónica: no se puede expresar la información contenida en su estructura sin un mecanismo de decodificación.

El desarrollo de un mecanismo de decodificación tal fue uno de los acontecimientos cruciales en el origen de la vida, hace alrededor de cuatro mil millones de años. Gradualmente fue evolucionando un sistema bioquímico fundamental en el que la secuencia de nucleótidos del ADN se transcribe primero a una secuencia complementaria de ARN mensajero (que se abrevia ARNm). Seguidamente, el ARN mensajero dirige el ensamblaje de aminoácidos en la secuencia lineal característica de una determinada proteína, proceso que recibe el nombre de traducción.

En la traducción hay otro tipo de ARN que desempeña un papel central: el ARN de transferencia (ARNt). Las moléculas de ARN de transferencia constituyen un grupo de pequeñas cadenas globulares de polinucleótidos (a diferencia de las cadenas fibrosas de polinucleótidos como las del ADN o el ARNm), de una longitud de 75 a 90 nucleótidos. Actúan de vehículo en la transferencia, a la cadena ensamblada de la proteína, de los aminoácidos que se encuentran libres en la célula. Esta función vital, de intermediario entre el lenguaje de los ácidos nucleicos del código genético y el lenguaje de los aminoácidos de la célula activa, ha hecho del ARN de transferencia un objeto principal de la investigación en la biología molecular. Recientemente se ha logrado determinar, con gran exactitud, la estructura tridimensional de la molécula de ARNt, lo que supone un importante paso hacia la comprensión del proceso de la traducción en términos moleculares precisos.

La traducción de la secuencia de nucleótidos del ARN mensajero en proteína comprende dos fases principales. Primero, mediante una reacción catalizada por un enzima de gran tamaño llamado aminoacil-ARNt sintetasa, un aminoácido se une a una determinada molécula de ARN de transferencia. Hay muchos tipos diferentes de sintetasa en las células vivas, cada uno de los cuales especifica uno de los 20 aminoácidos distintos que se encuentran en las proteínas. Por ejemplo, la leucil-ARNt sintetasa une consigo selectivamente al aminoácido leucina y al ARNt de la leucina, lo que da lugar a la formación de un complejo de leucina y ARNt de la leucina, que queda libre. Una vez que un ARNt ha unido a sí a un aminoácido, está ya preparado para participar en la segunda fase de la síntesis proteica.

Esta segunda fase, del engarce de aminoácidos en una cadena, se lleva a cabo en el interior de un orgánulo celular llamado ribosoma, que es un agregado de más de 50 moléculas proteicas distintas junto con tres moléculas de ARN. El ribosoma constituye una intrincada pieza de la maquinaria molecular, destinada a facilitar la traducción de la secuencia de nucleótidos del ARN mensajero en secuencia polipeptídica de proteína. Aunque no se han desentrañado todos los pormenores de este proceso, sí se conocen los rasgos generales del mismo.

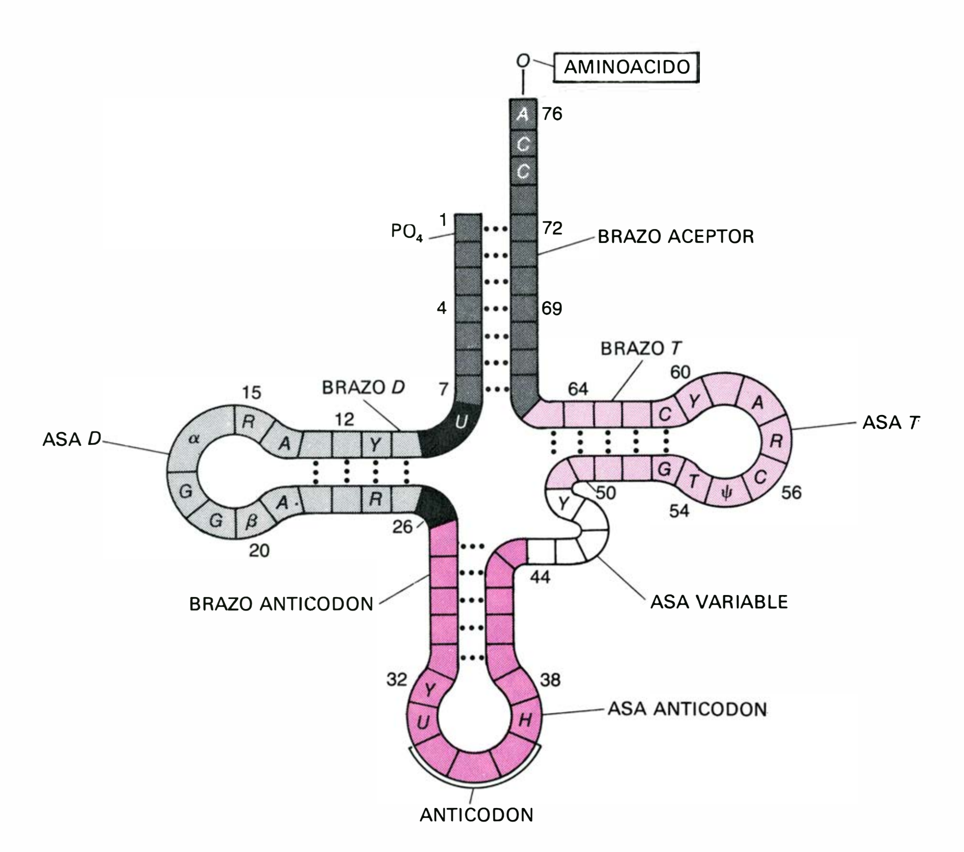
Cada aminoácido de una proteína está especificado, en el filamento del ARN mensajero, por un grupo de tres bases nucleotídicas adyacentes que recibe el nombre de codón. Hay cuatro tipos de bases nucleotídicas en el ARN mensajero; por tanto, existe un total de 4^3, o 64, codones posibles. La relación que media entre los codones y los aminoácidos que aquéllos especifican es el código genético. El hecho de que el código sea, al parecer, el mismo en todos los organismos vivientes constituye una prueba notable de la unidad de la vida a nivel molecular.

En los ribosomas hay dos zonas que están involucradas en la traducción. Una de ellas es el *locus A*, considerado el puente de unión del aminoacil-ARNt. En esta posición es por donde la molécula de ARN de transferencia y el aminoácido unido a ella se enganchan al ribosoma. El ARNt se aloja allí en parte debido a un conjunto de interacciones específicas con el ARN mensajero, que se había asociado poco antes al ribosoma. Tres bases nucleotídicas especiales del ARN de transferencia, denominadas anticodón, interactúan con las tres bases complementarias del codón del ARN mensajero. En la interacción están implicados los enlaces débiles dirigidos que se conocen por puentes de hidrógeno, en los que un átomo de hidrógeno con una pequeña carga positiva es compartido por otros dos átomos con

una pequeña carga negativa. Los puentes de hidrógeno constituyen también la fuerza que mantiene unidas las bases nucleotídicas complementarias de la doble hélice de ADN: la base guanina de un filamento de la hélice se aparea siempre con la base citosina del otro filamento, y la base adenina se aparea siempre con la base timina.

Adyacente al *locus A* del ribosoma se encuentra el lado de unión del peptidil-ARNt, o *locus P*. La molécula de ARN de transferencia, con la cadena de aminoácidos en crecimiento unida a ella, está enlazada a este *locus* e interactúa específicamente con el siguiente triplete de bases de la cadena de ARN mensajero. En el transcurso de la síntesis de la proteína, la cadena polipeptídica en crecimiento es desenganchada de la molécula de ARNt del *locus P* y transferida al extremo del solitario aminoácido unido a la molécula de ARNt del *locus A*. Una vez que la transferencia ha tenido lugar (las reacciones de separación y acoplamiento las lleva a cabo un enzima ribosómico), la cadena polipeptídica en crecimiento se ha alargado en un aminoácido. Se libera luego, del *locus P*, la molécula "vacía" de ARNt, y el ARN mensajero, junto con el peptidil-ARNt alargado, son desplazados del *locus A* al *P*. Un nuevo ARN de transferencia con un aminoácido unido a él se dirige ahora al ribosoma y se acomoda en el *locus A* desocupado, mediante interacciones específicas entre las bases de su anticodón y las del siguiente codón del filamento de ARN mensajero. El sistema ha retornado a su punto de partida, dispuesto a comenzar otro ciclo de procesos en el cual otro aminoácido más quedará añadido a la cadena. Esta adición escalonada va repitiéndose hasta que se haya sintetizado la proteína en su totalidad.

El proceso de elongación de la cadena polipeptídica es bastante rápido; tiene lugar hasta veinte veces por segundo en una célula bacteriana y alrededor de una vez por segundo en una célula de mamífero. La molécula de hemoglobina, por ejemplo, es una proteína grande, que consta de cuatro cadenas polipeptídicas de unos 140 aminoácidos cada una. La síntesis de una de estas cadenas tardaría siete segundos en una célula bacteriana y de dos a tres minutos en una célula de mamífero. Aún siendo esta tasa de síntesis bastante elevada hay, sorprendentemente, pocos errores en la traducción, ya que la maquinaria del ribosoma asegura un ajuste muy preciso entre cada molécula de ARN de transferencia y el ARN mensaiero. El

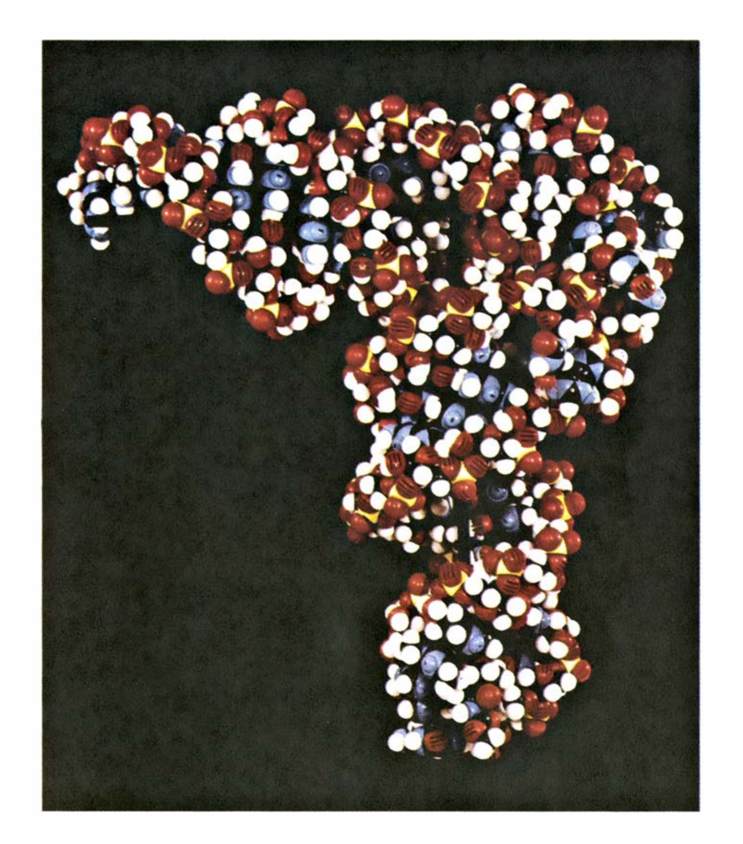
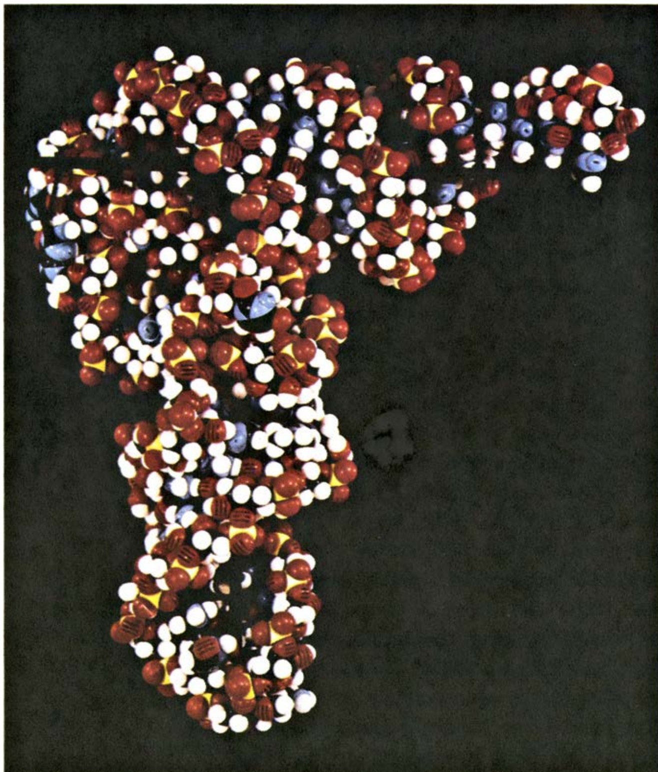


DISPOSICION EN HOJA DE TREBOL. Tal es el modelo de plegamiento bidimensional de la molécula de ARN de transferencia (ARNt), deducida por primera vez en 1965, a partir de la secuencia de los nucleótidos componentes del ARNt de alanina de levadura. Desde entonces, se ha comprobado que esta disposición encaja con la secuencia de nucleótidos de unos 100 ARNt, aislados a partir de células animales, vegetales y bacterianas. Se indican las bases nucleotídicas que en todas las secuencias de ARNt se encuentran en la misma posición. Los brazos escaleriformes están compuestos por bases complementarias de distintas partes de la cadena de nucleótidos, que se aparean mediante enlaces de hidrógeno, provocando el plegamiento de la cadena sobre sí misma. El número de nucleótidos de los distintos brazos y asas es generalmente constante, salvo en las dos partes del asa *D* indicadas por α y β (que constan de uno, dos o tres nucleótidos en los diferentes ARNt), y en el brazo extra (que normalmente posee cuatro o cinco nucleótidos, pero puede llegar a tener hasta 21). Las abreviaturas corresponden a las siguientes bases nitrogenadas: *A* (adenosina), *G* (guanosina), *C* (citidina), *U* (uridina), *R* (adenosina o guanosina), *Y* (citidina o uridina), *T* (ribotimidina), Ψ (pseudouridina), *H* (adenosina o guanosina modificadas).

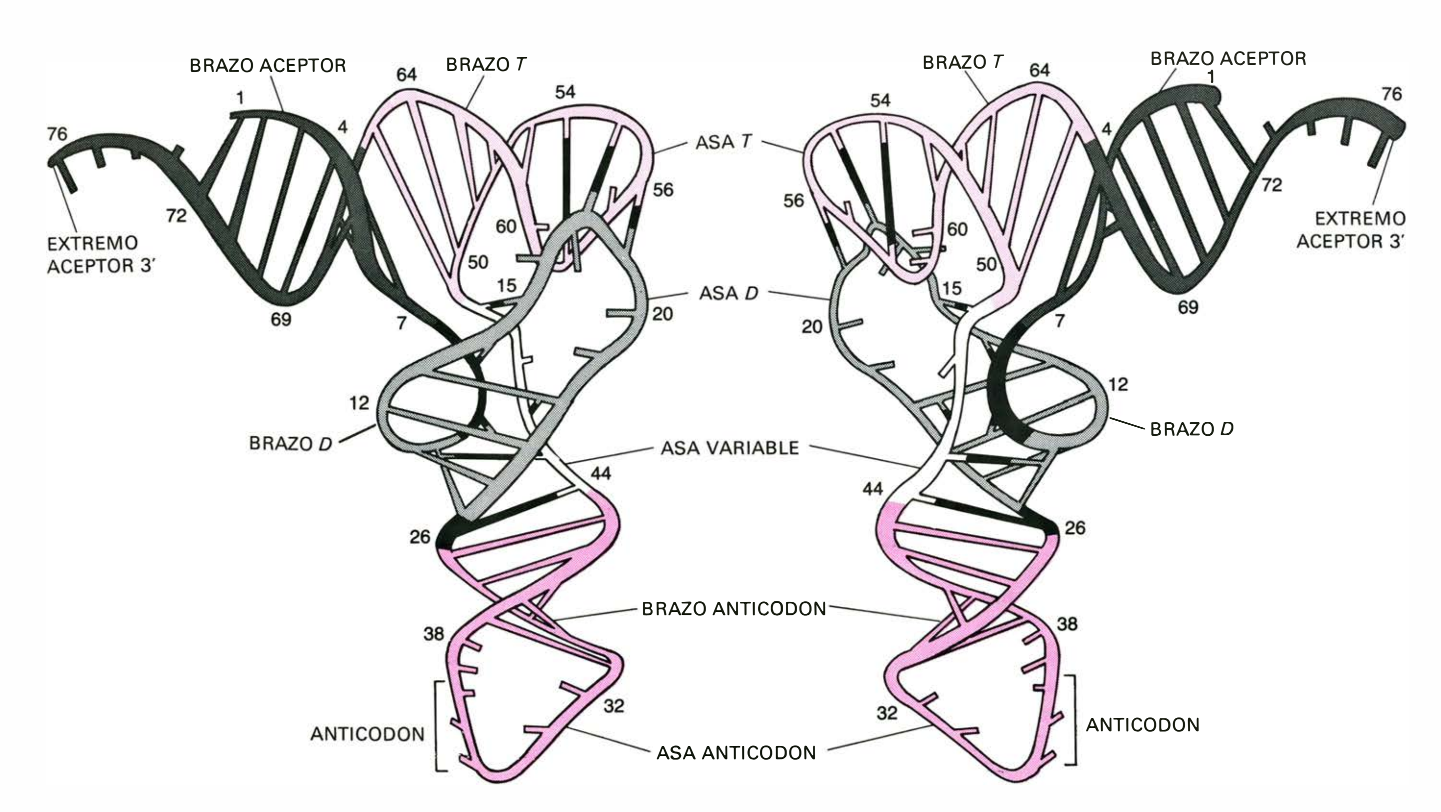
proceso es asimismo muy eficiente, dado que, por lo general, existen varios ribosomas traduciendo un único filamento de ARN mensajero.

A fin de comprender cómo el ARN de transferencia transporta un aminoácido al ribosoma y lo transfiere a la cadena polipeptídica en crecimiento, es esencial conocer la estructura tridimensional de la molécula de ARNt. Uno de los primeros indicios que aludían a dicha estructura surgió a partir de la secuencia de nucleótidos de un ARNt de levadura, específico para el aminoácido alanina, determinada en 1965 por Robert W. Holley y sus colegas en la Universidad de Cornell. Estos investigadores advirtieron que había ciertas regiones de la secuencia que resultarían complementarias en caso de que la cadena se plegara sobre sí misma. En particular, tales regiones podían formar enlaces de hidrógeno entre sí, de forma análoga al apareamiento de bases de la doble hélice de ADN (salvo que en el ARN la base adenina se aparea con uracilo en vez de hacerlo con timina). La cadena polinucleótida del ARN de transferencia podría, por tanto, disponerse de tal forma que contara con unas regiones, denominadas brazos, formadas por un doble filamento sostenido por puentes de hidrógeno, y otras, no apareadas, llamadas bucles o asas. La combinación propuesta de brazos y asas se asemeja a un trébol de cuatro hojas, por lo que se la conoce por disposición en hoja de trébol.

Característica de la secuencia de nucleótidos del ARN de transferencia es que, en ella, se incluyen algunas bases no usuales, siendo la mayoría de ellas bases normales de ARN que han sido modificadas por la adición de uno o más grupos metilo (CH_3). Debido a tal peculiaridad, algunas partes de la disposición en hoja de trébol han tomado el nombre

MODELO ESPACIAL del ARNt de la fenilalanina de levadura, que da una aproximación de la forma real de la molécula. Se construyó basándose en los análisis por difracción de rayos X realizados en el laboratorio del autor en el Instituto de Tecnología de Massachusetts y la Duke University School of Medicine. La cadena de nucleótidos del ARNt se pliega en una estructura compacta con forma de *L*. Durante la síntesis de proteínas, el aminoácido fenilalanina se une al extremo del brazo horizontal de la *L*. Tres bases nucleotídicas del extremo del brazo vertical reconocen el código genético para la fenilalanina en el filamento del ARN mensajero (ARNm). El aminoácido pasa a la cadena de proteína en crecimiento. Aquí, el carbono figura en negro, el oxígeno en rojo, el nitrógeno en azul, el fósforo en amarillo y el hidrógeno en blanco.



MODELO DE PLEGAMIENTO de la cadena de nucleótidos del ARN de transferencia de la fenilalanina de levadura. El esqueleto de fosfato y azúcar de la molécula se representa como un tubo enroscado indicando los peldaños transversales el apareamiento entre bases de los brazos. Los peldaños cortos designan bases no involucradas en enlaces de hidrógeno. El sombreado remite a la disposición en trébol de la página anterior.

de las bases modificadas que presentan. Por ejemplo, el asa *T* se denomina de esta forma porque incluye timina *(T)*, que se encuentra en el ADN, pero no en los distintos tipos de ARN, a excepción del ARN de transferencia. Análogamente, el asa *D* contiene frecuentemente la base modificada dihidrouracilo *(D)*. Otras regiones de la hoja de trébol son el brazo extra ("variable loop"), que tiene un número de bases distinto según el ARNt de que se trate (desde 4 hasta 21), el asa anticodón, en la que se hallan las tres bases del anticodón, y el brazo aceptor, que acepta el aminoácido específico de cada ARNt.

Un rasgo interesante del esquema en hoja de trébol es la presencia de secuencias de nucleótidos iguales en las 100 secuencias de ARNt que han sido determinadas hasta el momento. También es constante el número de pares de los brazos: siete en el brazo aceptor o aminoácido, cinco en el brazo *T*, cinco en el brazo anticodón y tres o cuatro en el brazo *D*. Estas características se mantienen en las moléculas de ARNt de plantas, animales, bacterias y virus. Es más, el esquema de brazos, asas y nucleótidos constantes que se observa en la hoja de trébol parece tener la misma universalidad que el código genético. La explicación de esta constancia vendría, en su mayor parte, tras el posterior conocimiento de la estructura tridimensional del ARNt.

Actualmente, se suele determinar la estructura tridimensional de las macromoléculas biológicas mediante análisis de cristales moleculares por difracción de rayos X. Un cristal molecular es un conjunto de moléculas empaquetadas en una red tridimensional regular. Si se dirigen hacia el cristal rayos X de una longitud de onda comparable a la distancia entre los átomos, aquéllos se difractan o dispersan en distintas direcciones por las nubes electrónicas de los átomos del retículo cristalino. El modelo de difracción del cristal se detecta como un conjunto de impresiones en una película sensible a los rayos X, siendo el oscurecimiento de la emulsión proporcional a la intensidad de cada haz dispersado.

Este modelo aporta gran cantidad de información sobre la estructura del cristal. Por un lado, la amplitud de la onda dispersada por un átomo es proporcional al número de electrones de éste, de suerte que un átomo de carbono desviará los rayos X con una fuerza seis veces mayor que otro átomo de hidrógeno. En segundo lugar, las ondas dispersadas se recombinan en el interior del cristal; según estén en fase o desfasadas, se reforzarán o anularán mutuamente. La forma de interferirse las ondas dispersadas depende sólo de la disposición en que se hallen los átomos del cristal, por lo que cabe reconstruir la imagen de una molécula a partir de su modelo de difracción.

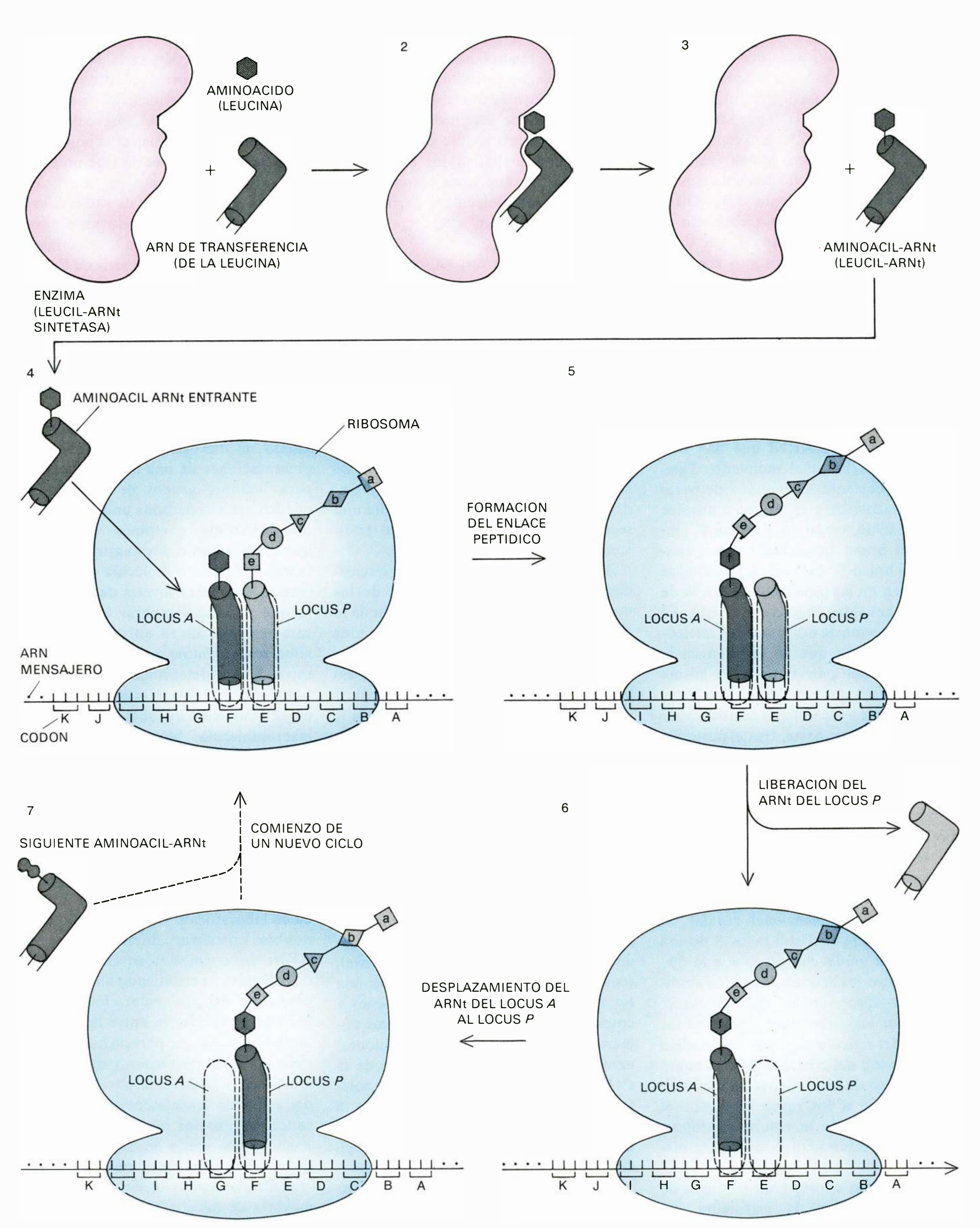
Para analizar la estructura tridimensional de una macroproteína o una molécula de ácido nucleico se prepara primero un cristal de la sustancia. Seguidamente, se monta el cristal en un tubo capilar, situándolo con una orientación precisa respecto al haz de rayos X y a la película. Se va girando el cristal alrededor de cada uno de sus ejes a fin de obtener series de fotografías de rayos X en las que aparezcan ordenaciones regulares de impresiones, de varias intensidades. Cada una de las fotografías representa una sección bidimensional de la red tridimensional de puntos.

A continuación, se miden las intensidades de todas las impresiones de los modelos de difracción, bien a partir de la película, bien mediante un contador Geiger. Se necesita, no obstante, una información adicional antes de poder establecer la estructura tridimensional, a saber, las fases de los haces de rayos X dispersados respecto de un punto arbitrario fijo en el cristal. Esta información se obtiene insertando como marcadores en la red cristalina diversos átomos de metales pesados, platino u oro por ejemplo. La adición de estos átomos cambia ligeramente el modelo de difracción y permite calcular las fases de los haces difractados.

Con esta información en la mano se puede deducir la densidad electrónica de un elevado número de puntos del cristal, separados regularmente, utilizando una serie de Fourier: una suma de senos y cosenos. Para el manejo del enorme número de términos (más de mil millones) involucrados en la determinación de la estructura de una proteína o de la molécula de un ácido nucleico, se precisa la ayuda de un ordenador de alta velocidad. La primera molécula de este tipo cuya estructura se determinó así fue la proteína mioglobina: la hazaña se realizó en 1958. Hoy en día, esta técnica se aplica casi rutinariamente en el análisis estructural de macromoléculas.

El producto final de la técnica es una representación tridimensional que muestra la distribución de los electrones en el cristal. La representación suele consistir en una serie de secciones paralelas superpuestas, siendo cada sección una lámina transparente de plástico en la que la densidad electrónica se indica por líneas negras de nivel, parecidas a las que vemos en los mapas topográficos. El factor limitante en la interpretación de las representaciones de la densidad electrónica es el grado de resolución de éstas, que está determinado por el número de intensidades de los haces difractados que se incorpora a la serie de Fourier. Por ejemplo, una representación de una resolución de seis angstroms, deducida a partir de las impresiones más internas del modelo de difracción, revela la forma general de la molécula, pero proporciona pocos detalles adicionales sobre su estructura. (Un angstrom son 10^{-10} metros, aproximadamente el diámetro de un átomo de hidrógeno.) Se necesitan representaciones de una mayor resolución para delinear grupos de átomos, que pueden estar separados unos tres o cuatro angstroms, o átomos individuales, que estén a uno o dos angstroms de distancia. Una macromolécula suele analizarse a diferentes niveles de resolución, lo que permite visualizar las diferentes características de su estructura. La máxima resolución de un análisis mediante rayos X está determinada, no obstante, por el grado de perfección del cristal. La mejor resolución que suele obtenerse con macromoléculas biológicas es de unos dos angstroms.

El primer paso del proceso –la cristalización de la molécula– se convirtió, en el caso del ARN de transferencia, en un gran obstáculo. En 1968, nuestro grupo del Instituto de Tecnología de Massachusetts, e investigadores de otros cinco laboratorios descubrieron que era posible cristalizar diferentes tipos de ARNt disolviéndolos en varias mezclas de solventes y permitiendo una lenta evaporación del solvente. Este avance produjo gran alegría entre los biólogos moleculares, ya que parecía quedar superado el mayor obstáculo, y que la determinación de la estructura tridimensional del ARN de transferencia estaba al alcance. A nuestro júbilo siguió rápidamente cierto grado de decepción, cuando se comprobó que, aun habiéndose cristalizado varios tipos de ARNt, la mayoría de los cristales eran totalmente desordenados. Debido a ello, los cristales ofrecían modelos de difracción de muy baja resolución (normalmente entre 10 y 20 angstroms) y, por tanto, apenas revelaban gran cosa sobre la estructura detallada de la molécula. Aunque nos alentaba el descubrimiento de que el



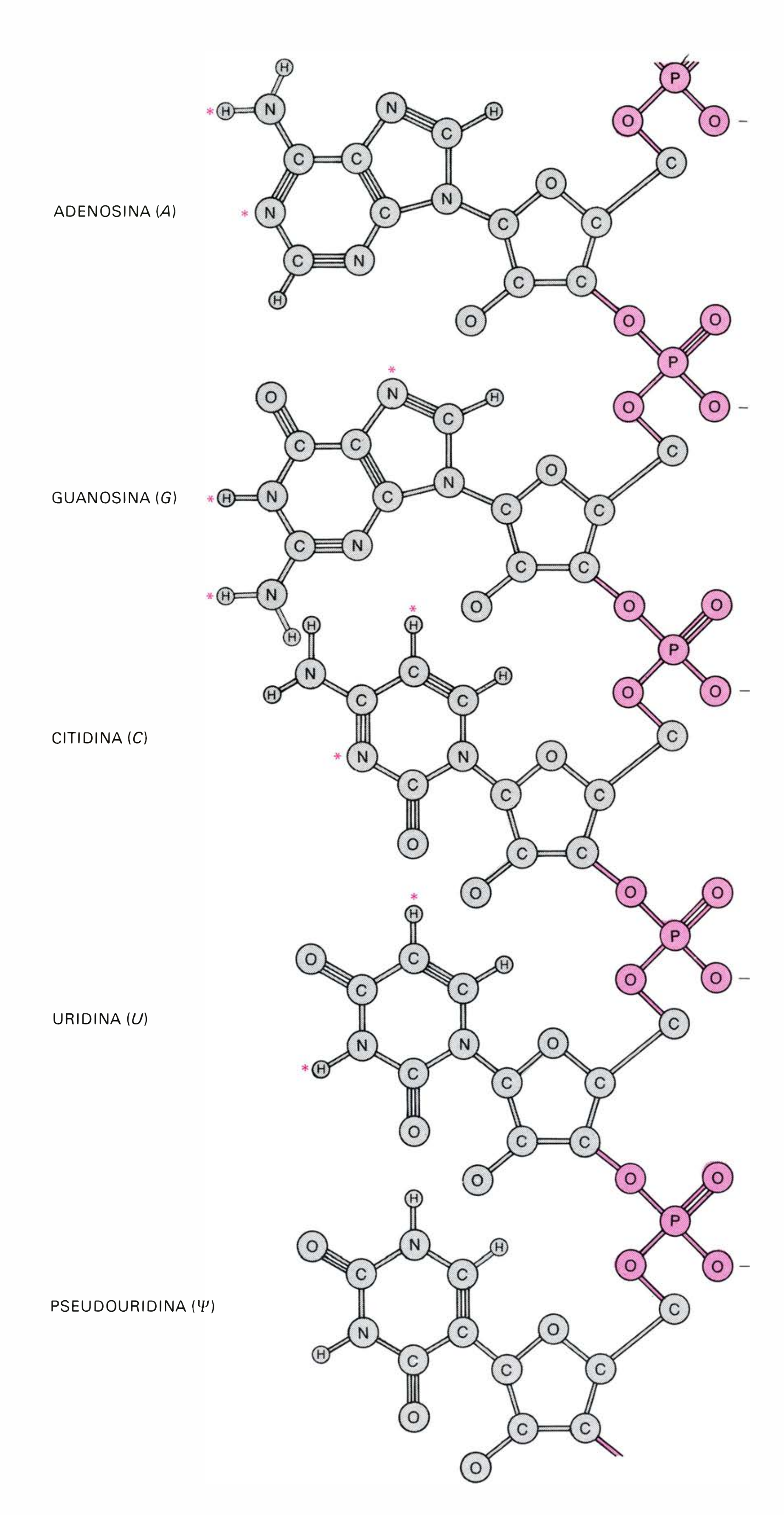
FUNCION DEL ARN DE TRANSFERENCIA, en la síntesis de una molécula de proteína, es hacer que la cadena de aminoácidos refleje la secuencia de nucleótidos del modelo representado por el ARN mensajero. Primero, un enzima de gran tamaño, denominado aminoacil-ARNt sintetasa, une un ARN de transferencia específico a su aminoácido correspondiente por medio de un enlace covalente *(1-3)*. El ARN de transferencia con el aminoácido unido a sí se acopla al *locus A* del ribosoma, que es el orgánulo en el que los aminoácidos se ensamblan para formar la cadena polipeptídica de una proteína. Este acoplamiento requiere unos enlaces de hidrógeno específicos entre las tres bases del codón del filamento de ARN mensajero, que especifican un aminoácido, y las tres bases del anticodón del ARN de transferencia *(4)*. Una molécula de ARN de transferencia situada en el *locus P* adyacente transfiere entonces la cadena polipeptídica en crecimiento al ARNt del *locus A (5)*. El ARNt "vacío" abandona el *locus P* y el ribosoma recorre la distancia de un codón a lo largo del ARN mensajero, de suerte que el ARN de transferencia que porta la cadena polipeptídica se ve desplazado del *locus A* al *P (6, 7)*. El ciclo de la síntesis comienza entonces de nuevo.

ARNt podía cristalizar, resultaba descorazonador comprobar que había que seguir trabajando para llegar a obtener un material adecuado para el análisis mediante difracción de rayos X.

Junto con Gary J. Quigley y Fred L. Suddath hicimos un esfuerzo conjunto para hallar las condiciones en que el ARNt diera lugar a un cristal bien ordenado, produciendo un modelo de difracción de rayos X con la suficiente resolución como para revelar la estructura tridimensional de la molécula. Durante dos años estudiamos un gran número de diferentes especies de ARNt y condiciones de cristalización. Finalmente, llegamos a un importante descubrimiento: la adición de espermina, una pequeña molécula cargada positivamente, daba lugar a la formación de cristales muy bien ordenados de un ARNt extraído de células de levadura, específico del aminoácido fenilalanina. El cristal estabilizado con espermina mostraba un modelo de difracción de una resolución próxima a los dos angstroms (un angstrom es igual a 10^{-10} metros).

A finales de 1972, trabajando con Alexander McPherson, Daryll Sneden, Jung-Ja Park Kim y Jon Weinzierl, obtuvimos una representación de las densidades electrónicas del cristal, en la que pudimos trazar el esqueleto de la cadena de nucleótidos del ARNt con una resolución de cuatro angstroms. Con esta resolución, no era posible distinguir individualmente las bases de la cadena, pero los grupos fosfato (PO_4), electrónicamente densos, podían verse como un rosario de cuentas enroscado tridimensionalmente. Sorprendentemente, descubrimos que la cadena de nucleótidos estaba dispuesta de tal manera que la molécula adquiría una conformación en *L*, estando formada una de sus ramas por el brazo aminoácido y el brazo *T* y la otra por el brazo *D* y el brazo anticodón. Las secuencias complementarias unidas por enlaces de hidrógeno, que habían sido identificadas en la disposición en hoja de trébol, se apreciaban claramente como dobles hélices de ARN. Las diferentes asas ocupaban posiciones estratégicas, bien en un extremo de la molécula, bien en el codo, donde las asas *T* y *D* se enroscaban de forma compleja.

Este plegamiento de la molécula era completamente inesperado. Durante los años anteriores, numerosos investigadores habían hallado las características comunes a las hojas de trébol de todos los ARN de transferencia e intentaron predecir cómo podría plegarse la molécula de ARNt. Como en tantas otras



LOS NUCLEOSIDOS, que constan de una base unida al azúcar ribosa, se enlazan mediante grupos fosfato (PO_4) cargados negativamente para formar la cadena polinucleotídica del ARN de transferencia. Los cuatro nucleósidos fundamentales de la molécula son la adenosina, la guanosina, la citidina y la uridina. El ARN de transferencia incorpora también muchos nucleósidos modificados, de los que se han identificado más de 50. La modificación más común es la sustitución de un átomo de hidrógeno por un grupo metilo (CH_3). Esta reacción está catalizada por enzimas especiales y tiene lugar en los puntos indicados por un asterisco. Hay otras modificaciones estructurales: la pseudouridina tiene la base unida a la ribosa por un átomo de carbono en vez de otro de nitrógeno.

ocasiones, la naturaleza demostró ser más sutil de lo imaginado. El plegamiento en forma de *L* explicaba muchas de las observaciones químicas que habían ido acumulándose, y produjo gran admiración comprobar los fines funcionales a los que esta forma tan poco usual se acomodaba.

A mediados de 1974, junto con Joel L. Sussman, Andrew H.-M. J. Wang y Nadrian C. Seeman, habíamos interpretado la representación de densidad electrónica con una resolución de tres angstroms. La forma general de la molécula era la misma que la que aparecía a cuatro angstroms, pero ahora se hacían visibles muchos más detalles, incluyendo la posición de la mayoría de las bases de nucleótidos. Casi al mismo tiempo, Jon Robertus, Brian F. C. Clark, Aaron Klug y sus colegas del British Medical Research Council Laboratory of Molecular Biology, de Cambridge, describieron su análisis cristalográfico con rayos X del ARN de transferencia, a una resolución de tres angstroms. Su ARNt era el mismo: el ARNt de la fenilalanina de levadura estabilizado con espermina, pero variaba la forma cristalográfica. Aun estando la molécula empaquetada de manera distinta en la red cristalina, la comparación de ambas estructuras tridimensionales resultantes de los análisis mostró que eran virtualmente idénticas. La concordancia entre los hallazgos de ambos grupos proporcionó una importante prueba en el sentido de que la estructura de la molécula de ARNt es independiente de cómo esté empaquetada en un cristal.

La representación de la molécula de ARNt con una resolución de tres angstroms confirmó nuestro hallazgo anterior de que está organizada en dos columnas de bases nucleotídicas dispuestas en ángulo recto. Estas columnas tienen regiones helicoidales y regiones no helicoidales, correspondiendo con los brazos y asas de la disposición en hoja de trébol. La representación de gran resolución reveló, además, que las dos regiones helicoidales comprenden unos 10 pares de bases, lo que corresponde a una vuelta de la doble hélice, y poseen el mismo tipo de enlaces de hidrógeno entre bases complementarias que el que se encuentra en la doble hélice de ADN.

MODELO DE DIFRACCION DE RAYOS X, uno de los patrones utilizados por los autores para deducir la estructura tridimensional del ARN de transferencia. El modelo se obtuvo dirigiendo rayos X a una red cristalina de moléculas de ARNt y recogiendo los haces dispersados en una película. Las impresiones (puntos) informan de la distribución de los electrones en el interior del cristal.

En las partes no helicoidales de la molécula de ARNt, muchas de las bases están orientadas con sus grupos formadores de enlaces de hidrógeno dirigidos hacia el interior de la molécula, participando en diferentes interacciones mediante enlaces de hidrógeno poco usuales, denominadas interacciones terciarias. Estos enlaces pueden darse entre dos o tres bases que habitualmente no se consideran complementarias, entre una base y el esqueleto de fosfato y ribosa de la cadena de ARN de transferencia, o incluso entre distintas partes del esqueleto mismo. Es de particular interés el hecho de que en varias interacciones terciarias del ARNt estén implicados los grupos hidroxilo (OH) de la ribosa, ya que los grupos hidroxilo no están presentes en las moléculas glucídicas del ADN. Estas interacciones terciarias no son necesarias en una cadena lineal regular como la del ADN, pero son esenciales para estabilizar el complejo arrollamiento de la cadena polinucleotídica del ARNt.

Se encontró un enlace de hidrógeno no usual en el brazo aminoácido o aceptor, donde se presentaba el par de bases guanina-uracilo en lugar del par guanina-citosina o adenina-uracilo, que son los normales. La posibilidad de un apareamiento semejante había sido sugerida algunos años antes, al hacer Francis H. C. Crick la observación de que, aparentemente, debían darse ciertos tipos adicionales de complementación en la posición de la tercera base durante la interacción del codón del ARN mensajero con el anticodón del ARN de transferencia. Una de las complementaciones "no convencionales" que había postulado Crick era la del par guanina-uracilo, que estaría sostenido por dos enlaces de hidrógeno, resultantes de un ligero desplazamiento lateral de una de las bases. Los continuos análisis y el perfeccionamiento de la representación de densidades electrónicas hasta una resolución de 2,5 angstroms confirmaron este tipo de apareamiento entre la guanina y el uracilo del brazo aminoácido.

Se han descubierto nuevas disposiciones de los enlaces de hidrógeno entre las interacciones entre bases de la molécula de ARN de transferencia *(véase la ilustración de la página 52).* La variedad de estas interacciones fue uno de los hallazgos más sorprendentes que se desprendieron de nuestro trabajo para la determinación de la estructura.

La mayoría de las bases de nucleótidos del ARNt están superpuestas en dos



MODELO DETALLADO DEL ESQUELETO del ARNt de la fenilalanina de levadura; muestra las interacciones mediante enlaces de hidrógeno entre las bases. Se configuró en 1974, a partir de un estudio cristalográfico mediante rayos X a una resolución de tres angstroms. La figura mostrada aquí fue generada mediante ordenador. El esqueleto de ribosa y fosfato de la molécula figura en color, en tanto que las bases en gris.

columnas que forman los brazos de la molécula en *L*. Esta disposición explica la especial estabilidad del ARNt. Si se calienta una solución de moléculas del ARN de transferencia, éstas se desnaturalizan, es decir, la cadena de nucleótidos se desenrolla y vuelve a adoptar una conformación originaria. Este comportamiento es muy distinto del que manifiestan la mayoría de proteínas, que se desnaturalizan irreversiblemente; la ovoalbúmina, por ejemplo, se vuelve blanca y opaca al hervir un huevo, y se mantiene así al enfriarlo.

¿Por qué vuelve la molécula de ARN de transferencia con tanta rapidez a su estructura nativa? Es sabido que la interacción que se origina entre las bases adyacentes del interior de la doble hélice de ADN por estar superpuestas, es uno de los factores que más estabilizan dicha molécula. Análogamente, las bases del ARNt son predominantemente hidrófobas (repelen el agua), de forma que se retraen del solvente circundante hacia el interior de la cadena plegada; este comportamiento ayuda la molécula de ARNt a volver a su conformación nativa, más estable. En las proteínas no suele haber una interacción parecida que haga replegar espontáneamente la cadena polipeptídica. Por tanto, parece que la estructura del ARNt está organizada para preservar la superposición de sus bases. Simultáneamente, cierta arquitectura molecular de gran complicación mantiene las dos columnas superpuestas formando un ángulo recto.

Un aspecto importante de las interacciones terciarias encontradas en el ARNt de la fenilalanina de levadura es que, en las secuencias de nucleótidos de todos los ARNt, están involucradas las mismas bases. Más aún, las bases que se presentan en las regiones de la cadena que tienen un número variable de nucleótidos no suelen hallarse superpuestas; se localizan en asas que sobresalen de la superficie de la molécula de ARNt. Estos hallazgos sugieren que el armazón estructural del ARNt de la fenilalanina de levadura puede ajustarse a las secuencias de nucleótidos encontrados en otros ARNt. Por ejemplo, una región variable del asa *D* del ARNt de la fenilalanina de levadura contiene dos nucleótidos, y este segmento de la cadena polinucleotídica se dobla por fuera de la molécula y retorna. Si hubiera más nucleótidos en esta zona puede suponerse que el abultamiento sería mayor; y viceversa, si hubiera menos nucleótidos, sería menor. El tamaño de estas asas variables, de todas formas, no afectaría al patrón general de plegamiento de la molécula.

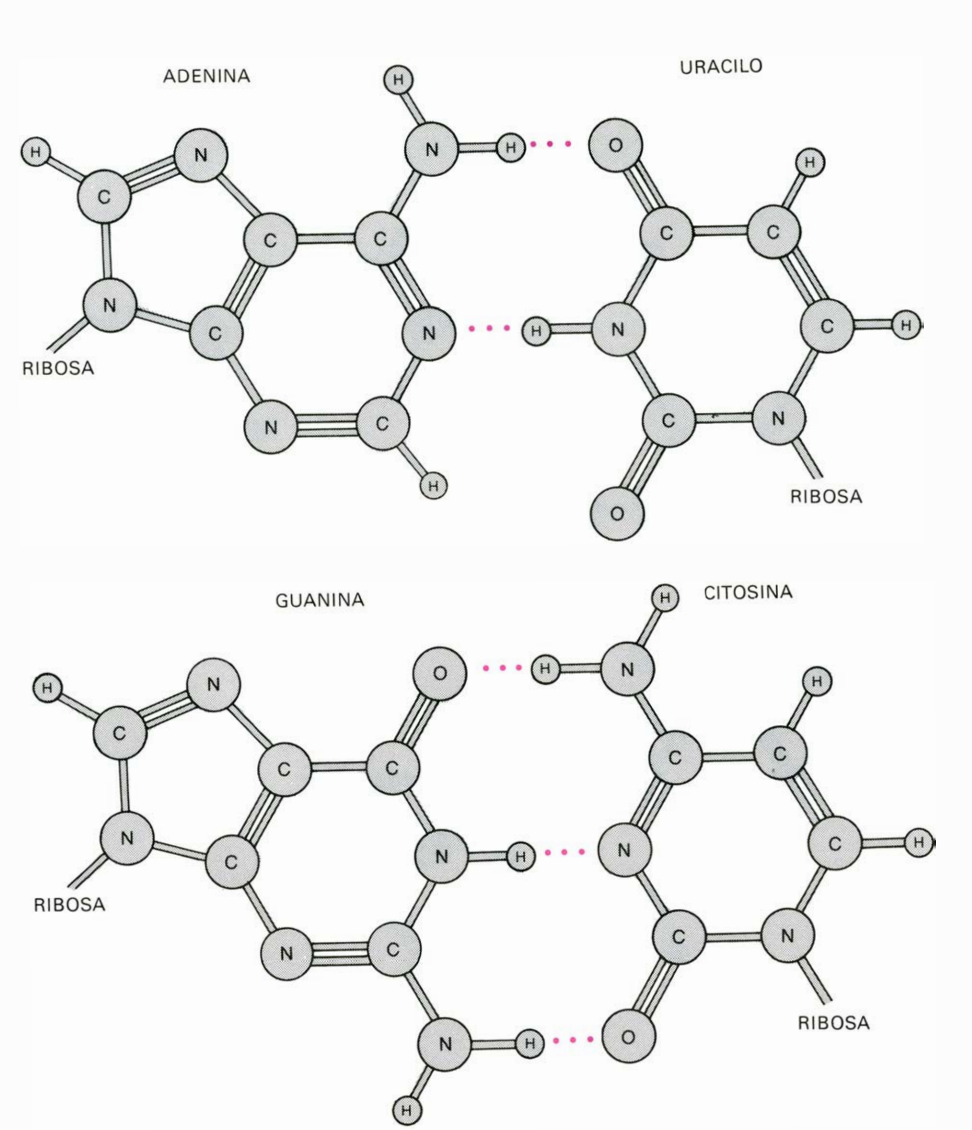
Muchos problemas importantes que conciernen a la estructura tridimensional de los ARN de transferencia permanecen, en general, sin resolverse. No está claro, por ejemplo, cuál debe ser la estructura detallada de los ARNt que tienen asas variables de gran tamaño. También es de interés la estructura de los ARNt "iniciadores", que dan comienzo a la síntesis de proteínas colocando el primer aminoácido. Algunos de los ARNt iniciadores tienen secuencias de nucleótidos que se apartan algo de las secuencias comunes de otros ARNt, particularmente en el asa *T*. Es muy verosímil que estas diferencias estén relacionadas con una estructura ligeramente distinta de la del ARNt de la fenilalanina de levadura.

Nuestros cristales de ARNt de fenilalanina de levadura contienen casi un 75 por ciento de agua. Conviene preguntarse si la molécula posee la misma forma en solución (donde es biológicamente activa) que cristalizada. Afortunadamente, ha habido numerosas investigaciones del ARNt de fenilalanina de levadura en solución. Los resultados obtenidos permiten correlacionar la estructura observada en el cristal con varias características químicas de la molécula. Por ejemplo, una de las peculiaridades del ARNt de la fenilalanina de levadura en solución es que algunos nucleótidos parecen prestarse rápidamente a cambios químicos al añadir reactivos a la solución, mientras que otros nucleótidos no. Esta disparidad resultaba desconcertante, hasta que se obtuvo la estructura de la molécula en el cristal. Quedó en-

LOS SEGMENTOS HELICOIDALES de la molécula de ARNt, correspondientes a los cuatro brazos de la disposición en hoja de trébol, se representan mediante bandas anchas en esta visión esquemática. Las dos regiones helicoidales se disponen en ángulo recto para proporcionar el armazón estructural al plegamiento en *L*. Cada región consta de unos 10 pares de bases, correspondiendo más o menos a una vuelta de la doble hélice. En estas regiones, la hélice es similar a la del ADN, salvo que en el ARNt los dos filamentos están formados por partes distintas de la misma cadena.

tonces claro que sólo ciertos nucleótidos, como los que se proyectan de la molécula en estado cristalino, pueden fácilmente sufrir modificaciones químicas. En general, se da perfecta correlación entre la sensibilidad de una región de la molécula de ARNt a modificaciones químicas y la accesibilidad de dicha región de la molécula en estado cristalino.

Otros experimentos llevados a cabo en disolución pueden interpretarse a la luz de la estructura tridimensional; entre ellos se cuentan los que se basan en la resonancia magnética nuclear, que es sensible a la estructura tridimensional de una molécula. Diversos investigadores han hallado una adecuada correlación entre las señales de resonancia magnética nuclear obtenidas a partir de moléculas de ARN de transferencia en disolución y la estructura deducida de los análisis por difracción de rayos X del ARNt de fenilalanina de levadura cristalizada. Estos y otros hallazgos proporcionan una prueba convincente de que la estructura de la molécula de ARNt cristalizado es la estructura de la forma biológicamente activa de la molécula.



LA COMPLEMENTACION mediante enlaces de hidrógeno entre las bases de las regiones helicoidales de la molécula de ARN de transferencia sigue el modelo propuesto por primera vez por James D. Watson y Francis H. C. Crick para la doble hélice de ADN (con la excepción de que en el ARN de transferencia el uracilo sustituye a la timina). La adenina y el uracilo se aparean por medio de dos enlaces de hidrógeno, mientras que la guanina y la citosina lo hacen mediante tres.

Estas correlaciones son importantes, ya que la razón principal por la que se determina la estructura de una biomolécula es la de comprender cómo funciona en un sistema biológico. ¿Qué nos ha comunicado la estructura del ARN de transferencia sobre su funcionamiento? Al respecto debe hablarse con mucho más prudencia, pues los experimentos necesarios aún no se han efectuado. Primero, nos agradaría saber cómo el enzima aminoacil-ARNt sintetasa reconoce y selecciona únicamente el ARNt correcto para unirlo a un aminoácido específico. Para llegar a comprender totalmente este proceso, será necesario determinar la estructura tridimensional de la sintetasa cuando forma el complejo con el ARNt, de modo que pueda apreciarse la naturaleza de las interacciones específicas entre el enzima y el ácido nucleico. En la actualidad, se están llevando a cabo estudios de este tipo en muchos laboratorios, y las respuestas deberán estar al caer en un próximo futuro. Ya sugieren algunos experimentos que, en el proceso de reconocimiento, ciertas regiones de la molécula de ARNt son más importantes que otras.

Otra cuestión fundamental es: ¿por qué la molécula de ARNt tiene forma de *L*, y se sitúa el anticodón a más de 76 angstroms de distancia del aminoácido? No se ha obtenido todavía la respuesta definitiva, si bien parece probable que la forma del ARNt esté relacionada con su función esencial de transferencia en el interior del ribosoma. Puede que, para la célula, haya sido necesario desarrollar una molécula de ARNt plegada de forma tan peculiar, a fin de que dos ARNt adyacentes en los loci *A* y *P* puedan disponerse conjuntamente sobre el filamento de ARN mensajero, de suerte que la cadena polipeptídica en crecimiento logre transferirse de uno a otro. Quizás el brazo aminoácido de la *L* gire en el interior del ribosoma para que la cadena de proteína pueda ser transferida al ARNt enlazado con el siguiente codón de la cadena de ARN mensajero. Este segundo ARNt uniría la cadena a su propio aminoácido y, a su vez, volvería a pasar la cadena. La considerable distancia que hay entre el final del brazo aminoácido y el anticodón de la molécula de ARNt puede ser también funcionalmente importante, en el sentido de que diversas proteínas ribosómicas puedan interactuar simultáneamente con varias regiones del ARNt, a fin de colaborar en el mantenimiento de la precisión de la síntesis proteica.

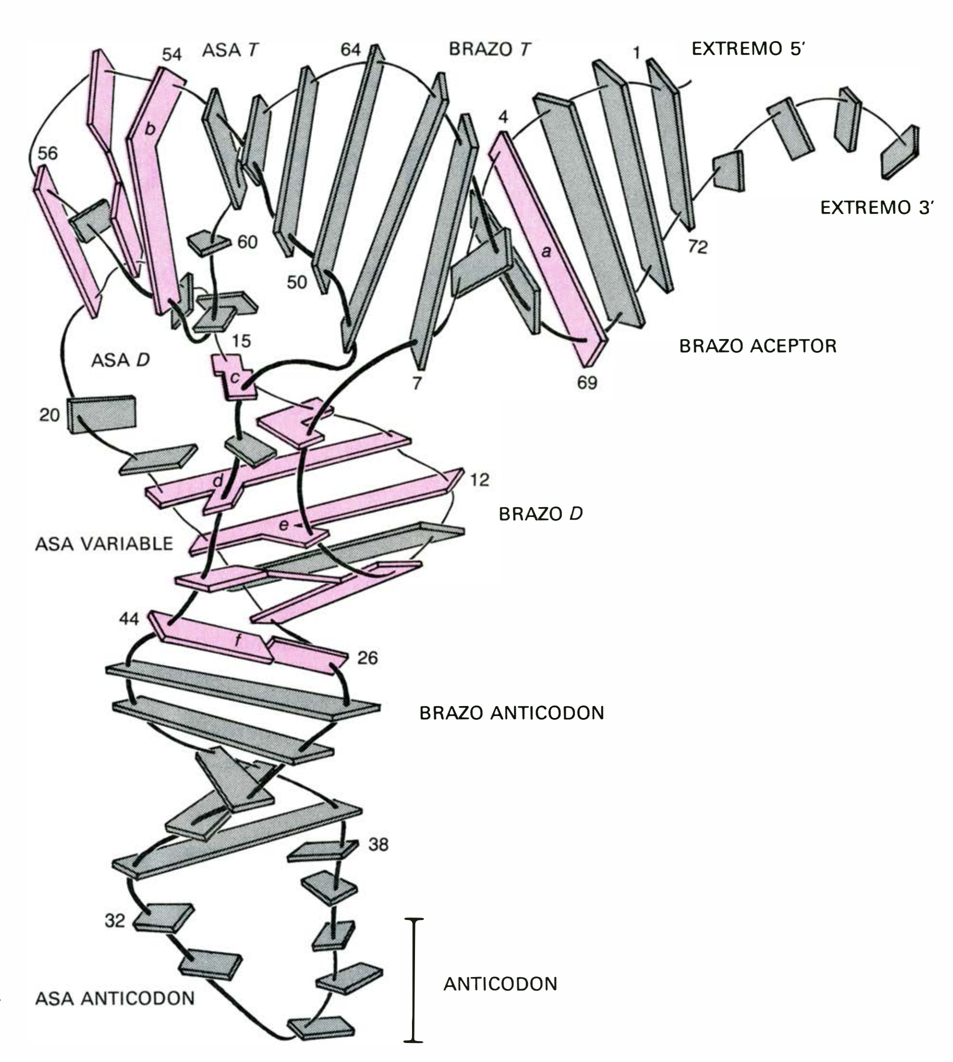
La configuración de la molécula de ARNt obtenida a partir del análisis por difracción de rayos X es esencialmente estática. En su entorno natural dentro de la célula, puede que la molécula sufra cambios conformacionales, particularmente al interactuar con estructuras macromoleculares, como los ribosomas. Experimentos recientes sugieren que, en el interior del ribosoma, cabe la posibilidad de que las asas *D* y *T* se separen al desplazarse la molécula del *locus A* al *P*. También es posible que se modifique la forma del asa anticodón al ponerse

LAS INTERACCIONES NO USUALES entre las bases estabilizan el modelo de plegamiento de la molécula de ARN de transferencia. En el brazo aminoácido o aceptor, las bases guanina y uracilo, que normalmente no son complementarias, se mantienen unidas por dos enlaces de hidrógeno como resultado de un ligero desplazamiento lateral de una de ellas *(a)*. En el asa *T*, la 1-metil-adenina se aparea con timina, una forma modificada de uracilo que posee un grupo metilo más que éste *(b)*. En el núcleo de la molécula, inmediatamente por debajo del codo, se aparean la guanina y la citosina, pero con dos enlaces de hidrógeno en vez de hacerlo con tres, que es lo normal. Este apareamiento es de tipo *trans*, ya que las ribosas quedan en posiciones opuestas *(c)*. También en la región del núcleo hay dos sistemas complejos de enlaces de hidrógeno que afectan a tres bases del mismo plano *(d, e)*. En la región donde se unen el brazo *D* y el brazo anticodón, una guanina dimetilada se aparea con una adenina por medio de dos enlaces de hidrógeno *(f)*. Este par de bases no queda incluido en un plano, ya que sobresalen los grupos metilo. Las bases están ladeadas 25 grados una de otra, como si fueran las aspas de una hélice. La guanina dimetilada está en la parte inferior del brazo *D* ("apilada" entre las otras bases) y la adenina en el extremo superior del brazo anticodón; ello estabiliza la unión de ambos brazos.

en contacto con el ARN mensajero dentro del ribosoma. Una evaluación más completa de estas proposiciones deberá aguardar los resultados de ulteriores investigaciones.

Al comenzar este artículo describimos con cierto detalle el papel del ARN de transferencia en la síntesis de proteínas, ya que es ésta la más esencial de las funciones moleculares en los sistemas biológicos. Sin la molécula de ARNt, la información genética no podría expresarse en la síntesis proteica. Adicionalmente, el ARNt participa en otros muchos procesos, algunos de los cuales son de gran importancia. Por ejemplo, las moléculas de ARNt pueden ceder aminoácidos a moléculas proteicas preformadas o a la estructura molecular de la pared celular bacteriana, independientemente del ribosoma.

Otro proceso en el que participa el ARNt es el control de la expresión génica. Se sabe de ciertos tipos de ARNt de transferencia portadores de un aminoácido que determinan si ciertos genes, es decir segmentos de ADN, se expresan o no, regulando su transcripción a ARN mensajero. No se conoce el mecanismo detalladamente, pero en algunos sistemas la función de control está asociada a un determinado nucleótido modificado de la molécula de ARNt, por ejemplo, un uracilo que se ha convertido en pseudouracilo. Se cree que el ARNt ayuda a controlar la expresión de muchos genes distintos, aunque no se conozca su número exacto. Los casos más intensamente estudiados han sido los de los genes que regulan la síntesis de aminoácidos, en donde la regulación debida al ARNt desempeña un papel fundamental.



ORIENTACION DE LAS BASES en la molécula de ARN de transferencia. El esqueleto de nucleótidos se ha reducido a una línea delgada. Las placas cortas representan bases no apareadas y, las mayores, pares de bases. Las letras hacen referencia a los diagramas moleculares de la ilustración anterior. Nótese la presencia de interacciones terciarias entre tres bases en la región nuclear de la molécula, por debajo del codo de la *L*. A grandes rasgos, la molécula está compuesta por dos columnas de bases nucleotídicas; ambas columnas se hallan superpuestas formando un ángulo recto entre sí. Las interacciones que son debidas a la superposición entre bases paralelas del interior de la molécula proporcionan una fuerza estabilizadora de fundamental importancia.

Otras observaciones sugieren que el ARN de transferencia puede estar implicado también en más tipos de regulación bioquímica. Por ejemplo, en el transcurso del desarrollo embrionario, una clase de modificación de ciertos nucleótidos de un ARNt da paso a otro tipo de modificaciones. De manera análoga, cuando una célula normal se vuelve cancerosa, los tipos de modificaciones de los nucleótidos de sus moléculas de ARNt cambian substancialmente. No se sabe todavía si estas transformaciones están asociadas con las funciones reguladoras del ARNt.

Otra área misteriosa concierne al relativamente elevado número de modificaciones de las secuencias del asa *D*, así como las del brazo extra. ¿Por qué la naturaleza se ha tomado la molestia de variar los nucleótidos que sobresalen de la superficie de la molécula? En general, se cree que estas secuencias no se requieren para la especificidad de la síntesis proteica; en cambio, tal vez se hallen involucradas en las funciones reguladoras de las moléculas de ARNt, ya que las regiones variables podrían proporcionar *loci* de reconocimiento específico para otras moléculas.

Finalmente, el ARNt está asociado a la síntesis de cadenas polipeptídicas y a la síntesis de cadenas de nucleótidos. Esta última síntesis la realizan enzimas especiales, como la transcriptasa inversa, descubierta hace unos años como constituyente de varios virus tumorales. La transcriptasa inversa sintetiza un filamento de ADN a partir de un molde de filamento único de ARN, lo que supone un sentido inverso al normal en la dirección de la información. Sorprendentemente, un tipo específico de ARNt se une primero al enzima y al ARN viral y señala el comienzo de la síntesis de la copia de ADN. Se desconoce completamente por qué se emplea para este fin un ARNt.

Es probable que moléculas semejantes al ARNt fueran un componente esencial de los primitivos sistemas vivientes. Una vez formadas estas moléculas, su estabilidad poco usual habría determinado que se utilizaran para fines distintos de aquel que es su función principal en la síntesis proteica. Aunque la elucidación de la estructura tridimensional del ARNt ha supuesto un importante paso adelante, resta aún mucho que aprender de esta versátil molécula y de sus múltiples funciones en la célula viva.

La estructura del medio interestelar

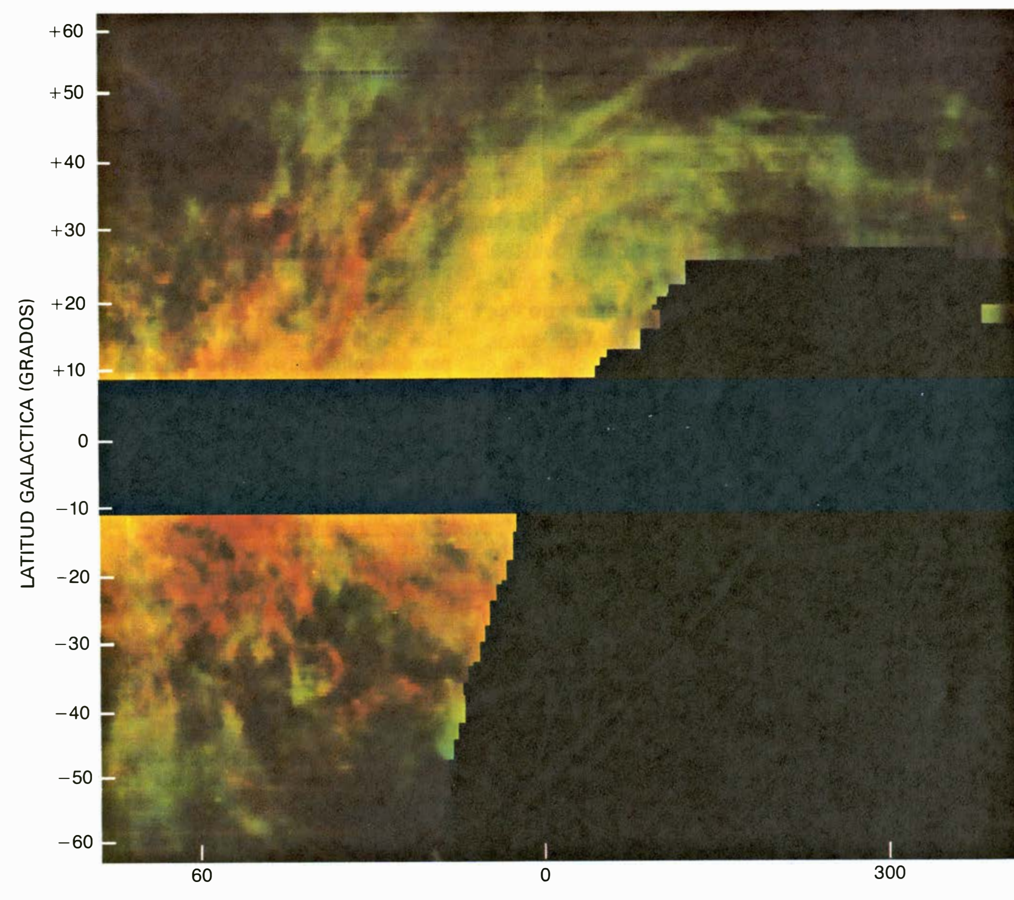
Se creía que el gas interestelar se distribuía probablemente en forma de nubes. Pero un análisis gráfico de las radioseñales emitidas por el gas pone de manifiesto que, en su mayor parte, se ordena en vastas envolturas

Carl Heiles

Hace medio siglo, las observaciones espectroscópicas de las estrellas lejanas revelaron que su luz se enrojecía debido a su propagación a través del gas y el polvo que impregnaban el espacio interestelar. El descubrimiento de la existencia de este material interestelar planteó una serie de interrogantes que en su mayoría permanecen sin respuesta. Una de las cuestiones fundamentales era la siguiente: ¿Cuál era la estructura del medio interestelar? En unas cuantas regiones del espacio, cerca de estrellas muy calientes, los átomos del gas interestelar se hallan ionizados, o han perdido uno o más de sus electrones, y resplandecen con una luz propia característica. Las primeras cuatro décadas de observaciones de tales regiones ionizadas sugirieron que el material interestelar no se distribuía uniformemente a través del espacio, sino que se agrupaba en enormes nubes más o menos esféricas esparcidas al azar a lo largo y ancho de toda la galaxia. Estas observaciones dieron lugar a un modelo teórico del medio interestelar conocido como el modelo de nubes. De acuerdo con este modelo, las nubes de gas interestelar presentaban un diámetro de unos 14 parsecs y una distancia promedio entre nubes de unos 30 parsecs. (Un parsec es 3,26 años luz.)

Sin embargo, el grueso del material interestelar se encuentra demasiado frío como para emitir energía en las longitudes de onda correspondientes al espectro visible y, consiguientemente, permanece invisible para los telescopios ópticos. A lo largo de la última década, los astrónomos han podido extender a otras regiones del espectro de radiación sus observaciones del medio interestelar, a las regiones correspondientes a la radiación X, ultravioleta, infrarroja y radio. Estas observaciones, en particular los mapas detallados de la distribución de hidrógeno recopilados para la longitud de radio-onda de 21 centímetros, han alterado radicalmente nuestros puntos de vista sobre la naturaleza y la estructura del gas interestelar.

Un átomo de hidrógeno consta de un protón y un electrón, cada uno de los cuales gira en torno a su propio eje. Los espines pueden ser paralelos o antiparalelos entre sí. Si son paralelos, el electrón puede invertirse de modo que su espín sea antiparalelo con respecto al espín del protón. Cuando el espín del electrón se invierte, el átomo emite radiación con



FOTOGRAFIA SINTETICA DEL MEDIO INTERESTELAR, que pone de manifiesto que el gas se dispone en largos filamentos curvados. La fotografía se sintetizó a partir de los datos obtenidos para una longitud de radioonda de 21 centímetros con el radiotelescopio de 26 metros del Hat Creek Radio Observatory. El autor y sus colaboradores estudiaron el medio interestelar en todo el cielo visible desde California, excluyendo una ancha banda centrada en el plano de la galaxia, debido a que en dicha zona la abundante cantidad de gas impide que las formaciones individuales pueden distinguirse. La gran parábola obscura de la izquierda representa la porción de cielo meridional invisible desde la latitud del observatorio; el pequeño rectángulo negro, próximo

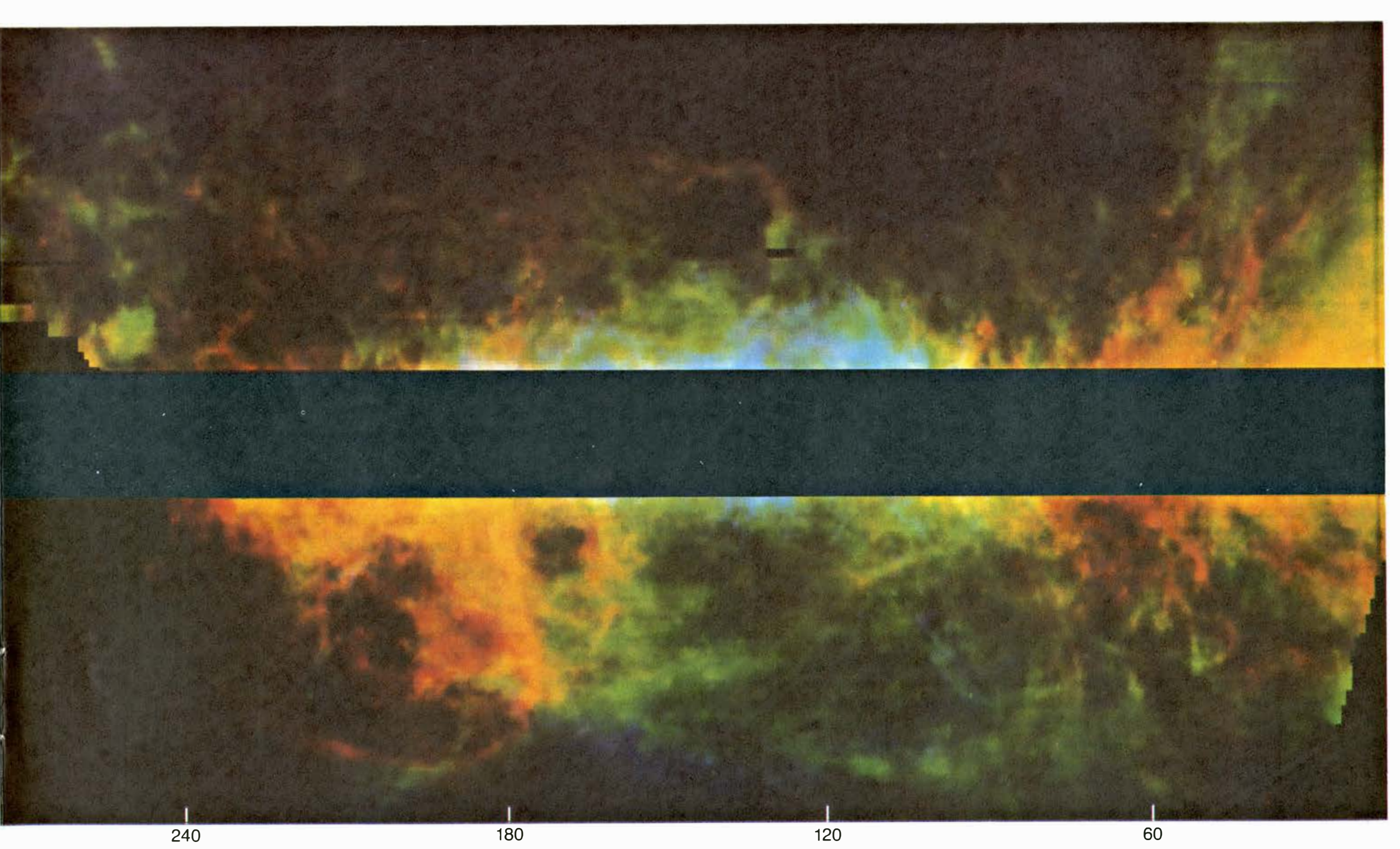
una frecuencia de 1420,4057 megahertzios, que corresponde a una longitud de onda de 21 centímetros, en la zona radio del espectro de radiación. El hidrógeno, que es el elemento más abundante del universo, constituye el 70 por ciento de la masa del gas interestelar. Más de la mitad del hidrógeno se encuentra en forma molecular (H_2) o está ionizado, es decir, sin su electrón. Sin embargo, en sólo una pequeña fracción del espacio interestelar se encuentra hidrógeno molecular e hidrógeno ionizado: allí donde el medio es relativamente denso. El hidrógeno esparcido por el resto del volumen espacial se halla en forma de átomos individuales.

Las observaciones del hidrógeno atómico para longitudes de radio-ondas presentan una ventaja decisiva con respecto a las realizadas para las longitudes de onda visibles. En este último caso, se estudia el hidrógeno interestelar mediante la observación de cómo el gas absorbe la luz de las estrellas; por tanto, las observaciones del gas para estas longitudes de onda se ven limitadas a aquellas regiones del cielo en las que detrás del gas existe una estrella apropiada. Para longitudes de radio-ondas, el hidrógeno interestelar se estudia observando la radiación emitida por el propio gas: no se necesita tener como fondo una fuente de radiación. Por consiguiente, se pueden realizar observaciones para longitudes de radio-ondas en cualquier región y nos pueden revelar la distribución del gas en todo el cielo.

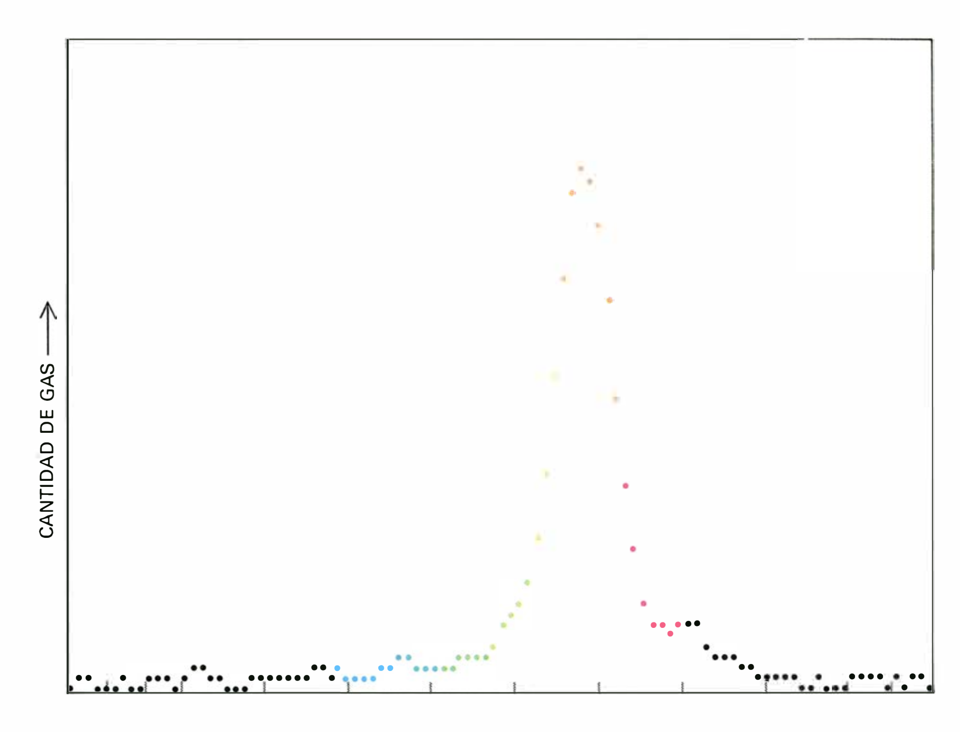
Primeras exploraciones celestes

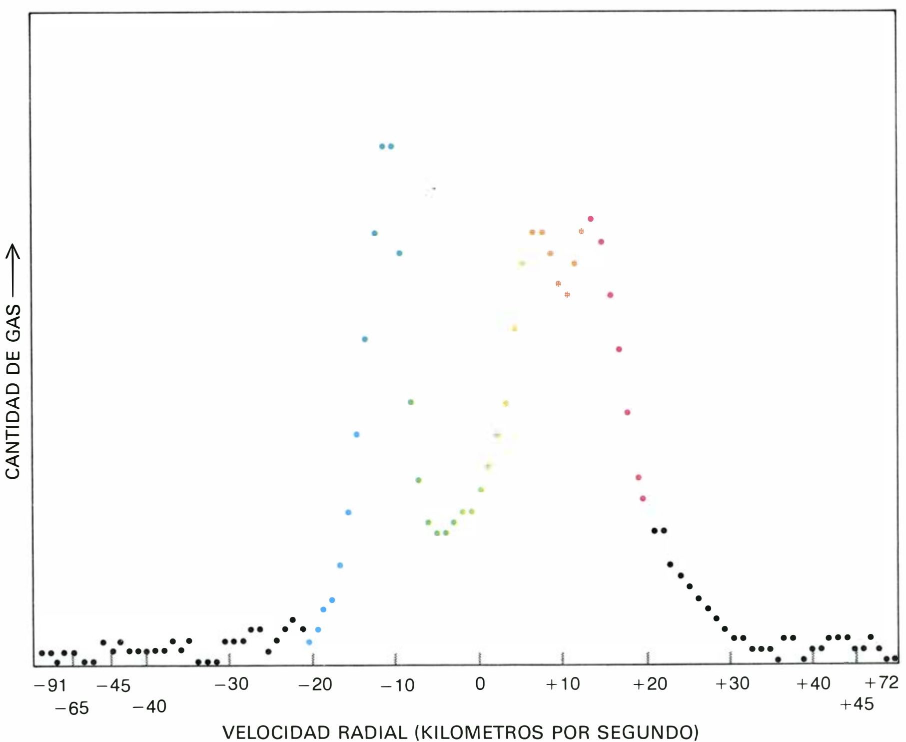
Los primeros intentos por realizar un mapa de la distribución de hidrógeno se limitaron a observaciones del gas en el plano central de la galaxia. Pronto se hizo patente que existe tal cantidad de gas en el plano galáctico que el gas de cualquier estructura no podría diferenciarse del resto del gas contenido en la dirección de observación. Posteriores exámenes se centraron esencialmente en regiones exteriores al plano galáctico, construyéndose mapas de zonas del cielo con un diámetro entre 10 y 20 grados. De acuerdo con el modelo de nubes del medio interestelar, en una región del cielo con un diámetro angular entre 10 y 20 grados deberían distinguirse docenas y quizá cientos de nubes de gas. Sin embargo, las primeras observaciones del medio interestelar exterior al plano galáctico revelaron un hecho sorprendente: las estructuras gaseosas registradas presentaban un tamaño angular tan extenso que se prolongaban más allá de los límites de las zonas estudiadas. Aunque parecían existir también formaciones de menor escala, se hallaban contenidas dentro de las estructuras mayores. Quedaba claro que no se comprenderían las formaciones menores hasta que no lo fueran las grandes.

Al no conocerse la extensión angular total de las estructuras mayores, resulta obvia la necesidad de realizar mapas de zonas del cielo más vastas. Es más, la astronomía es una ciencia estadística por su propia naturaleza: ya que los as-



a la longitud galáctica de 120 grados y a la latitud galáctica de + 30 grados, es el polo norte celeste, que no puede observarse por el telescopio habida cuenta de las características de su diseño. Adviértase que se repite una parte de la longitud galáctica, a fin de que pueda verse la continuidad de las formaciones estructurales a través de los límites de la gráfica. En la fotografía sintética el brillo de cada filamento representa la cantidad de gas en esa región. El color de cada formación representa la velocidad del gas y la dirección en que se está moviendo. Hay una progresión regular del color a través del espectro con respecto a la velocidad: el rojo representa gas que se aleja del sistema solar a 20 kilómetros por segundo, el amarillo-verde representa gas que no posee velocidad radial, y el azul representa gas que se aproxima hacia el sistema solar a 20 kilómetros por segundo. Cerca del plano de la galaxia puede observarse que, en general, el color del gas varía del azul al rojo según la longitud galáctica. La variación del color se produce por el hecho de que la galaxia no gira como un disco sólido.





PERFILES DE VELOCIDAD del medio interestelar medidos con el radiotelescopio de Hat Creek para 140.000 posiciones en el cielo. Cada perfil es una gráfica de la cantidad de gas que se mueve a diferentes velocidades radiales discretas. Se combinaron todos los perfiles para obtener las fotografías sintéticas. Las velocidades positivas indican que el gas se aleja del sistema solar; las velocidades negativas indican que el gas se aproxima. Cada punto representa una medición de la cantidad de gas en los 100 estrechos intervalos de velocidad considerados. Al haber, en el medio interestelar, muy poco gas que se mueva con altas velocidades con respecto al sistema solar, para velocidades superiores a ±45 kilómetros por segundo la anchura del intervalo de velocidad se ha tomado mayor. El aspecto del perfil de velocidad del gas interestelar es único para cada posición en el cielo. Al construir la fotografía sintética en color se imaginó que la longitud de onda de la luz visible aumentaba de izquierda a derecha a lo largo del perfil de radio-velocidades, con azul puro a menos 20 kilómetros por segundo y rojo puro a más 20 kilómetros. El perfil de velocidades superior se ha medido para una pequeña región próxima a la longitud galáctica de 181,7 grados y una latitud galáctica de −30,4 grados. En esta posición, la mayor parte del gas se aleja del sistema solar prácticamente a la misma velocidad. El perfil de velocidades es estrecho y vendrá representado por un color relativamente puro; en la fotografía sintética de las páginas precedentes la región es naranja rojizo. El perfil de velocidad inferior se obtuvo de una pequeña región en la longitud galáctica de 205,7 grados y la latitud galáctica de −22,6 grados. En esta posición del cielo, parte del gas se aproxima al sistema solar *(pico de la izquierda)* y otra parte se aleja de él *(picos de la derecha)*. En la fotografía de las páginas precedentes la región presenta un color rojo claro.

trónomos no pueden modificar las condiciones existentes en el espacio con el objeto de someter a experimentación el comportamiento de los sistemas celestes, deben contentarse con observar los objetos celestes en la multiplicidad de condiciones que ofrece la naturaleza. Para observar una muestra de las estructuras grandes del medio interestelar que sea estadísticamente significativa, debemos disponer de mapas del cielo realmente extensos. La decisión tomada por mis compañeros de la Universidad de California en Berkeley y por mí mismo consistió en observar todo el cielo, o, para ser más exactos, toda la parte del cielo que podía explorarse desde el Hat Creek Radio Observatory de la Universidad, a unos 274 kilómetros al norte de Sacramento, cerca del Lassen Volcanic National Park.

Con un radiotelescopio de 26 metros de diámetro realizamos un sondeo del cielo visible desde Hat Creek, excluyendo una banda de 20 grados de diámetro centrada en el plano galáctico donde se tiene tal cantidad de gas que las estructuras individuales se hacen indistinguibles. Para cubrir tal zona del cielo fue necesario orientar el telescopio hacia unas 140.000 posiciones distintas. La sensibilidad de las observaciones aumenta conforme mayor sea el tiempo dedicado a cada posición; empleamos unos dos minutos por posición con el fin de obtener los datos a la sensibilidad requerida. Las observaciones consumieron siete meses de tiempo de telescopio, que viene a suponer un intervalo de dos años.

Una multitud de nuevos datos

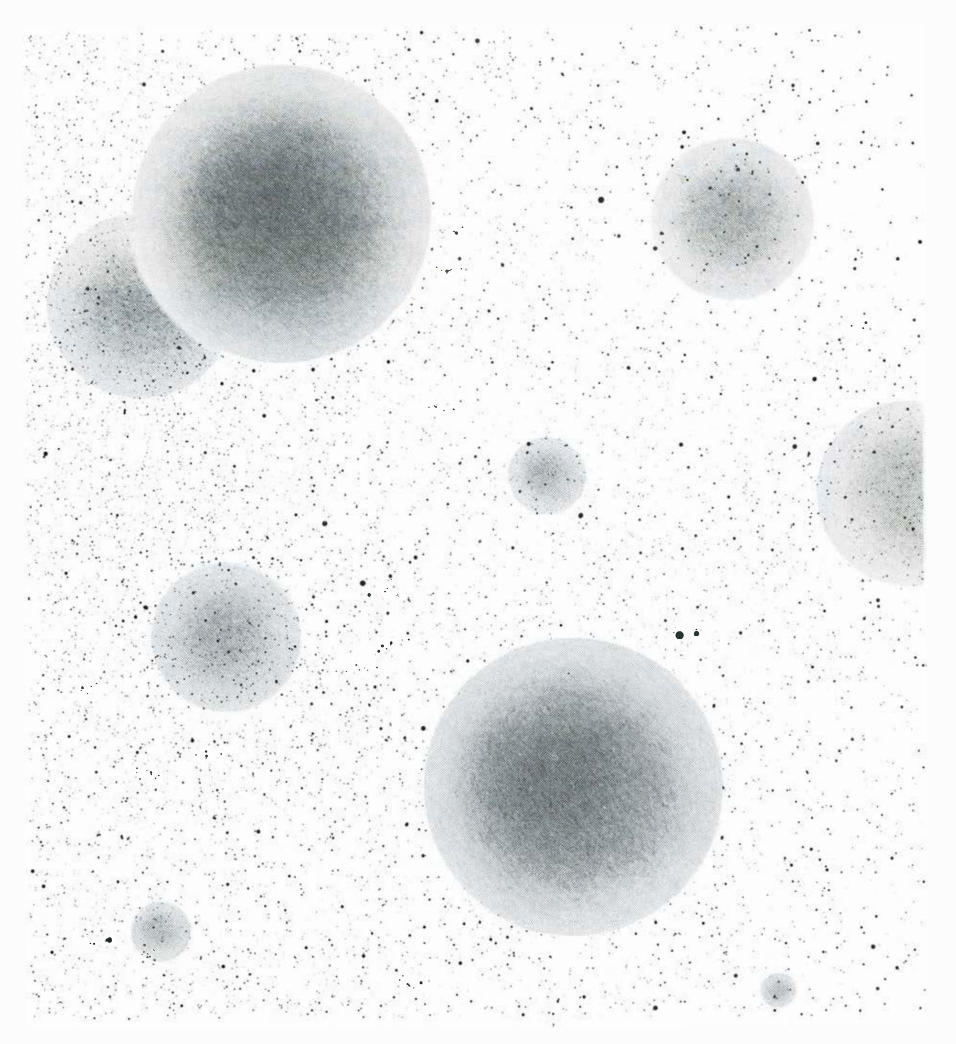
En cada posición recogimos información de la cantidad de gas y de su estado de movimiento. Aunque cada átomo de hidrógeno emite energía a una frecuencia de 1420,4057 megahertzios, la frecuencia a la que se recibe la radiación en el telescopio sufre un desplazamiento Doppler debido al movimiento del gas con respecto al observador. El desplazamiento que experimenta la frecuencia depende sólo de la dirección del movimiento y de la velocidad radial del gas: la frecuencia de la radiación emitida por el gas que se aleja del observador experimenta un corrimiento hacia las bajas frecuencias, mientras que si se mueve en dirección hacia el observador la variación se efectúa hacia las altas frecuencias. Por convenio, se atribuye velocidad positiva a una fuente que se aleje del sistema solar mientras que si se aproxima su velocidad es negativa.

Medimos la intensidad de la radiación del gas en 100 frecuencias distintas correspondientes a otros tantos intervalos de velocidad diferente.

La cantidad de gas que se mueve con una velocidad dada es proporcional a la intensidad de la radiación del gas que se mueve a esa velocidad. En muchos casos se presentan sobre la dirección de observación varias estructuras gaseosas estando cada una de ellas a diferente distancia del sistema solar. En consecuencia, cada estructura gaseosa presentará un corrimiento Doppler diferente y se observará en intervalos de velocidad distintos.

Un estudio de todo el cielo genera cantidades enormes de datos que plantean en sí mismas un serio problema. En nuestro caso, el problema era realmente grave. En efecto, puesto que habíamos medido las cantidades de gas correspondientes a 100 intervalos de velocidad en cada una de las 140.000 posiciones celestes consideradas, obtuvimos un total de 14 millones de datos puntuales de un espacio tridimensional en longitud galáctica, latitud galáctica y velocidad. Es un auténtico problema organizar tal cantidad de puntos de forma que puedan interpretarse adecuadamente por el cerebro humano. Tradicionalmente, los datos se presentan en radioastronomía mediante la elaboración de un mapa en el que la cantidad de gas viene representada por líneas de intensidad constante, de la misma forma que en los mapas topográficos la altura de las montañas se representa por medio de líneas de altitud constante. En un principio calculamos el mapa a partir de todos los datos obtenidos, pero actuando así resultaba que las estructuras a gran escala que queríamos estudiar aparecían difuminadas por un dédalo enmarañado de líneas de contorno correspondientes a las formaciones de menor tamaño contenidas en ellas. Resultaba evidente que había que echar mano de un método mejor para presentar la información obtenida.

Al observar el cielo a longitudes de onda visibles, los astrónomos obtienen información sobre grandes objetos que incluyen estructuras de menor escala, de manera casi rutinaria, mediante el registro de tales objetos sobre placas fotográficas. Mis compañeros y yo mismo necesitábamos disponer, pues, de una máquina que pudiese sintetizar una "fotografía" a partir de los datos digitales que habíamos conseguido y almacenado en una cinta magnética. Por fortuna, aunque para otros propósitos, tal máquina había sido desarrollada en el departamento de astronomía de la Princeton University. Además, varios astrónomos de Princeton estaban trabajando sobre la estructura del medio interestelar. Surgió en consecuencia, de forma natural, una estrecha colaboración.



MODELO DE NUBES de la estructura del medio interestelar; se propuso a partir de las observaciones para longitudes de onda visibles. El modelo propugnaba que el gas se agrupaba en nubes individuales *(gris)* con un radio de 7 parsecs y una distancia promedio internubes de 30 parsecs. (Un parsec es 3,26 años luz.) La distancia promedio entre estrellas *(puntos negros)* es de 0,4 parsecs.

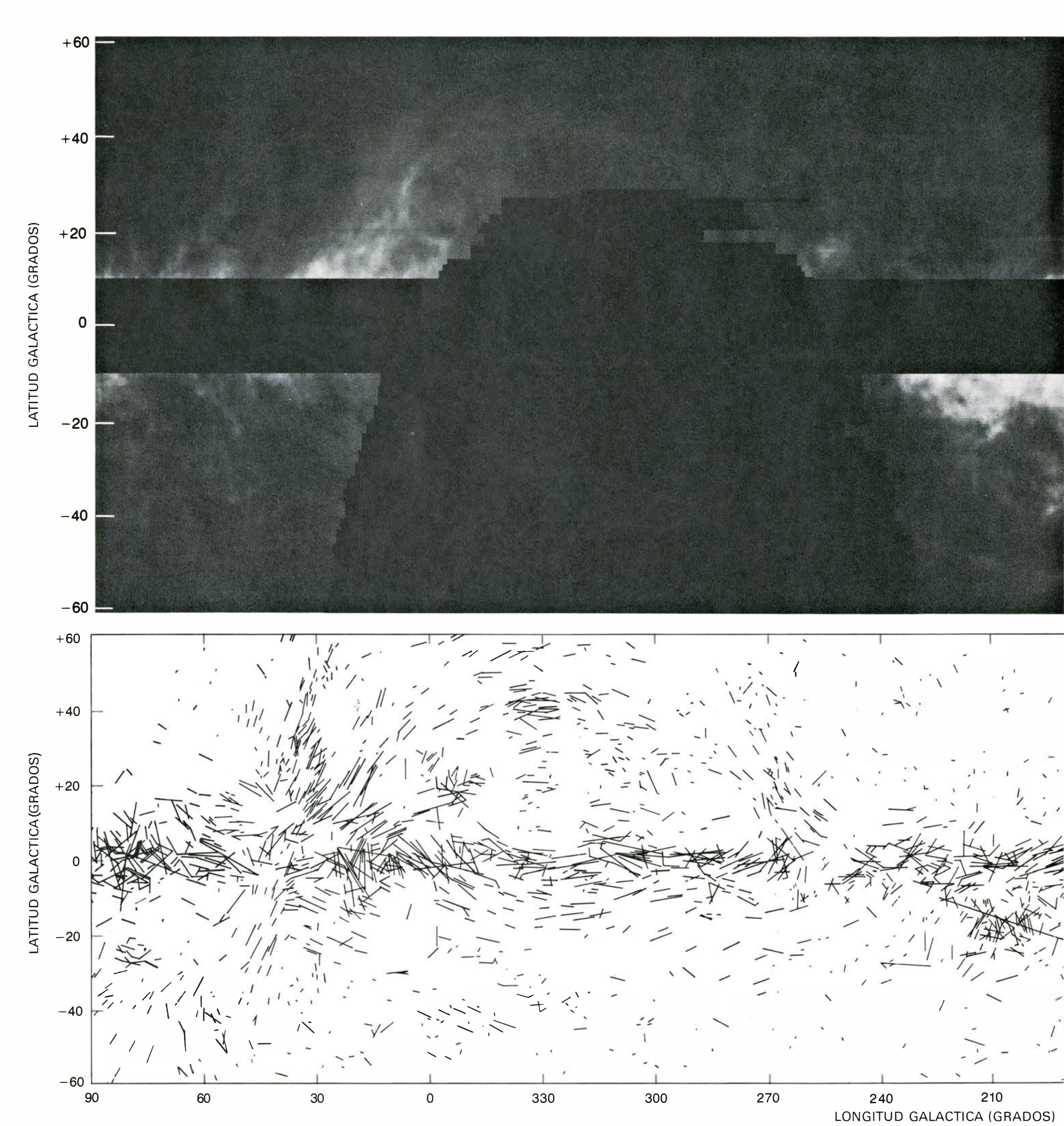
Síntesis de radiofotografías

Trabajando con Edward B. Jenkins, de Princeton, dejamos de lado, en una primera fase de la investigación, la información que poseíamos sobre los intervalos de velocidades individuales y nos limitamos a realizar una fotografía sintética de la cantidad de gas contenida en un amplio intervalo de velocidad, comprendido entre menos 20 kilómetros por segundo y más 20 kilómetros por segundo. Casi todo el gas se hallaba contenido dentro de este gran intervalo *(véase la ilustración de las dos páginas siguientes)*. El primer hecho que apreciamos en la fotografía sintética fue que el gas no tiende a organizarse en nubes individuales sino en largas estructuras filamentosas. Muchos de estos filamentos presentan una dimensión angular enorme extendiéndose algunos hasta ocupar un cuarto del diámetro del cielo observado. Pero no se puede determinar su longitud real en parsecs hasta conocer a qué distancia se encuentran del sistema solar.

Utilizando únicamente radiobservaciones del gas no podemos obtener información sobre dicha distancia. Necesitamos relacionar las observaciones del gas con las de estrellas, cuya distancia puede determinarse independientemente. La manera más eficaz y directa de relacionar ambos tipos de observaciones consiste en estudiar las líneas de absorción en el espectro de una estrella producidas por el gas interestelar que se encuentra entre la estrella y el sistema solar. Se producen las líneas de absorción debido a que el gas interestelar a baja temperatura absorbe las radiaciones, emitidas por la estrella más caliente que está detrás del mismo, correspon-

dientes a las mismas longitudes de onda de la radiación que sería emitida por el propio gas si se calentase hasta una temperatura lo suficientemente alta.

Jenkins y Blair D. Savage, de la Universidad de Wisconsin, han analizado las líneas de absorción en los espectros de las estrellas, creadas por el hidrógeno atómico interestelar. En razón de que tanto las líneas de absorción superpuestas sobre los espectros estelares como las líneas de radioemisión a 21 centímetros vienen ambas generadas por el hidrógeno atómico, una comparación entre las líneas de absorción y las de emisión revelará si todo el hidrógeno presente se encuentra localizado entre la estrella y nosotros o bien una parte de él se distribuye detrás de la estrella donde no afectará al espectro estelar. A continuación, determinando la distancia

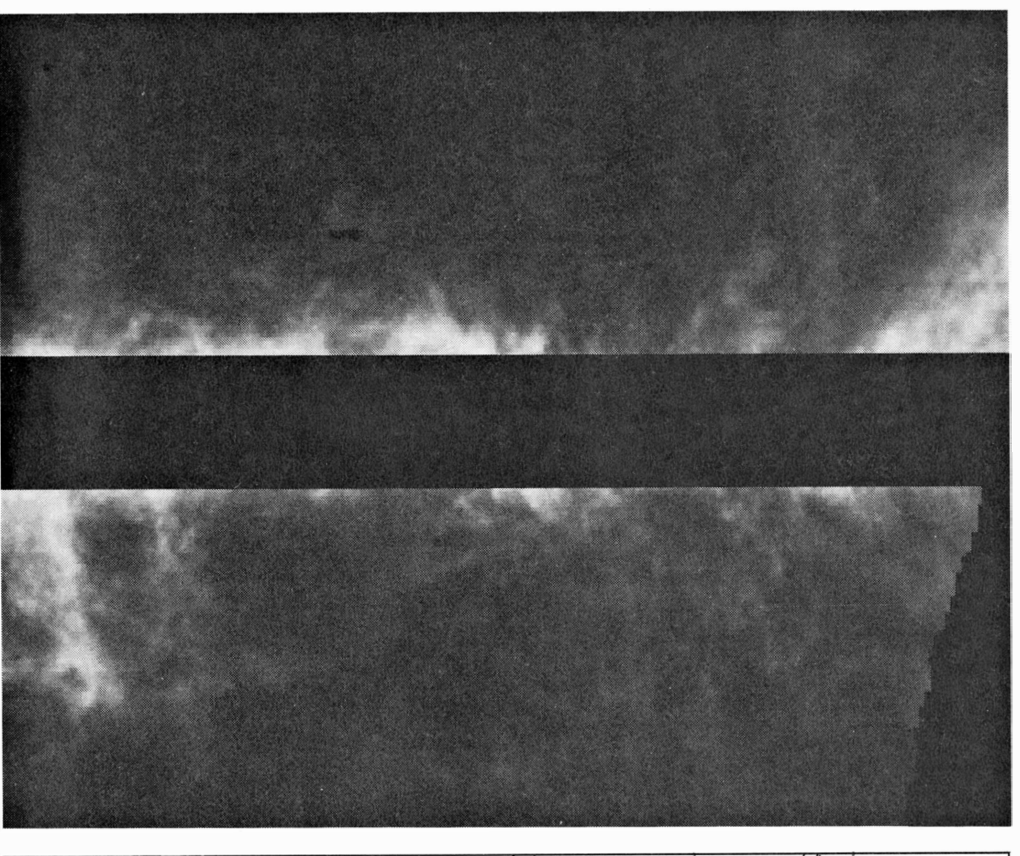


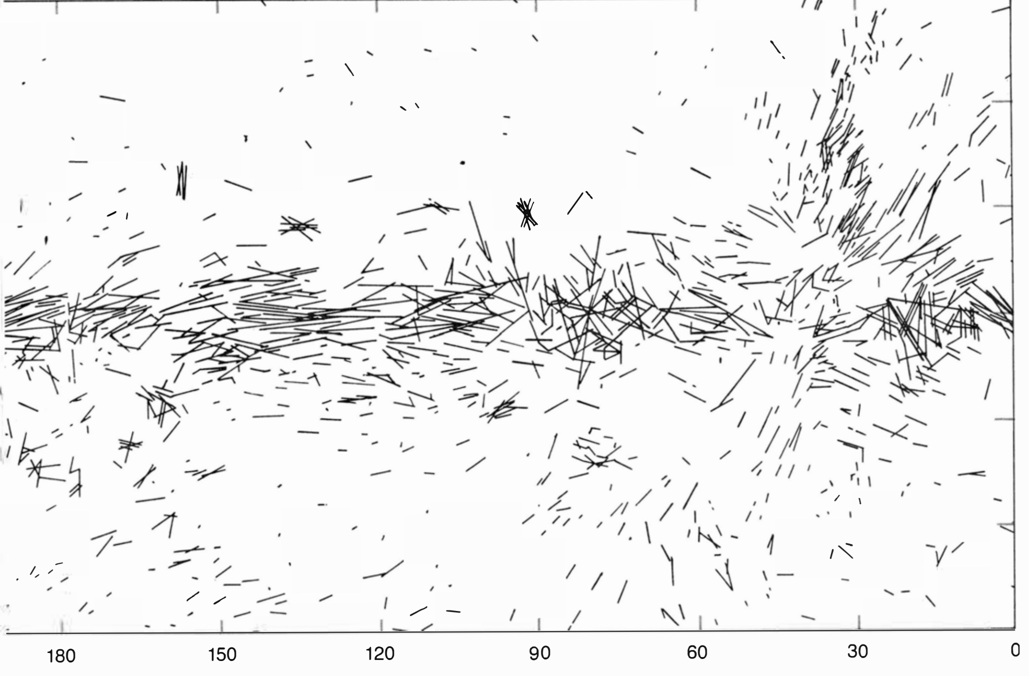
LOS FILAMENTOS SE ALINEAN con el campo magnético de la galaxia, como puede verse en esta comparación de una fotografía sintética del medio interestelar *(arriba)* con un mapa de polarización de la luz estelar *(abajo).* El polvo situado entre una estrella y el sistema solar polariza la luz de las estrellas de fondo. Cada línea representa una medición de la cantidad en la que la luz está polarizada; la longitud de la línea es proporcional a la intensidad de la polarización, y la dirección de la línea indica la dirección de la polarización. Arcos bien de-

de la estrella por medio de las técnicas astronómicas usuales, se puede tener la seguridad de que el gas que se sitúa delante de la estrella no estará más alejado del sistema solar de lo que lo está la propia estrella.

Las líneas de absorción producidas por el hidrógeno interestelar se localizan en la región ultravioleta del espectro, zona que no puede observarse desde la superficie de la tierra debido a la absorción atmosférica de este tipo de radiaciones. Sin embargo, disponemos de un telescopio, montado en el satélite *Copérnico* en órbita en torno a la tierra, que puede observar las longitudes de onda apropiadas, aunque presenta el inconveniente de que sólo es capaz de estudiar estrellas calientes que emitan intensamente en el ultravioleta. Tales estrellas no son frecuentes, pero las que existen proporcionan una información crucial que de otra forma no sería posible conseguir.

Existe una dilatada región del cielo con un tamaño angular de casi 30 grados en la constelación de Orion, que incluye también a la Gran Nebulosa de Orion, poblada con un buen número de estrellas calientes. Dicha región se encuentra próxima a la longitud galáctica de 210 grados y a la latitud galáctica de –20 grados. Las estrellas de esta región se hallan a una distancia de unos 500 parsecs del sistema solar. Los espectros ultravioleta de las estrellas de esta zona indican que hay poco gas delante de ellas. Las observaciones en el dominio radio muestran, sin embargo, que en esa misma parte del cielo hay mucho gas interestelar, el cual se integra en una gran formación estructural. Así pues, estos dos tipos de observaciones ponen de manifiesto que existe una zona del cielo bastante grande en la que hay poco gas a distancias del sistema solar inferiores a 500 parsecs. Estas observaciones contradicen la predicción obtenida a partir del modelo de nubes, que dice que la distancia media entre nubes de material interestelar es de 30 parsecs.





finidos del campo magnético en latitudes galácticas positivas, cerca de la longitud galáctica de 30 grados, se hallan notablemente alineados con los filamentos del gas que hay en su vecindad. Las zonas donde el campo se presente caótico y los filamentos no están alineados pueden ser regiones para las que existan varias concentraciones de gas entre la estrella y el sistema solar.

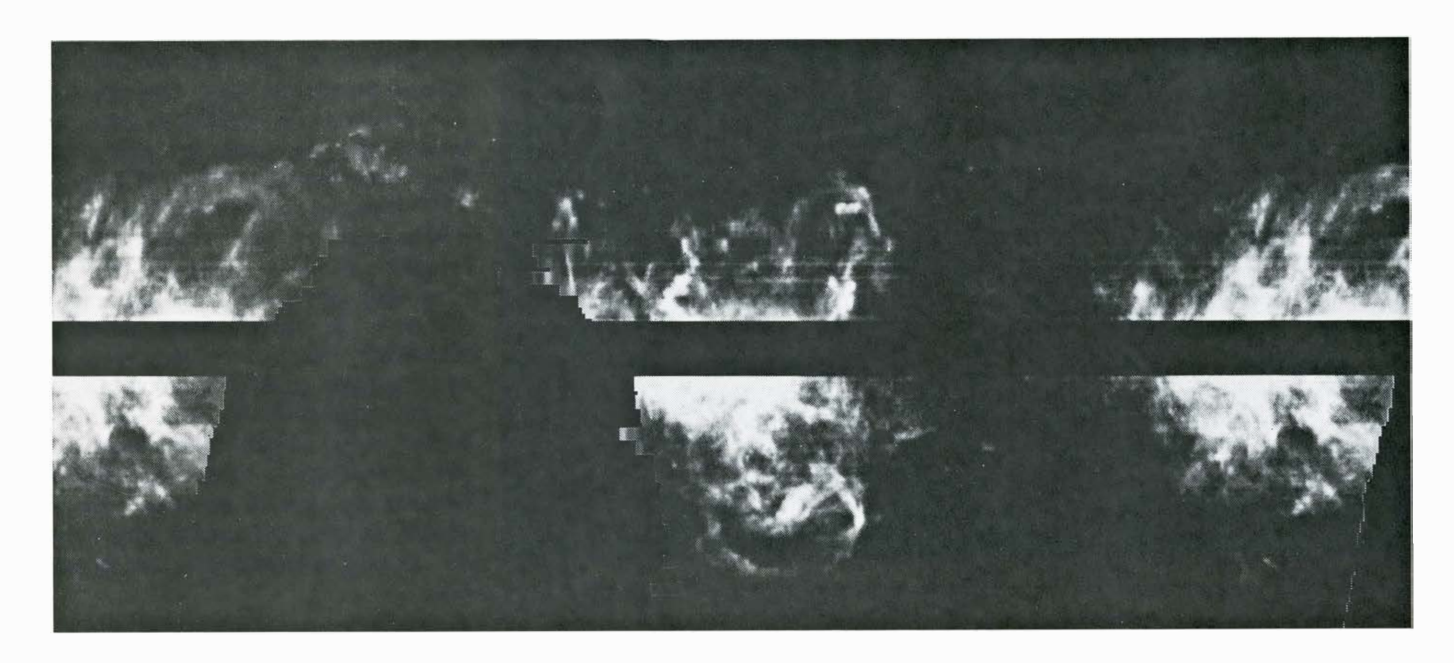
El campo magnético galáctico

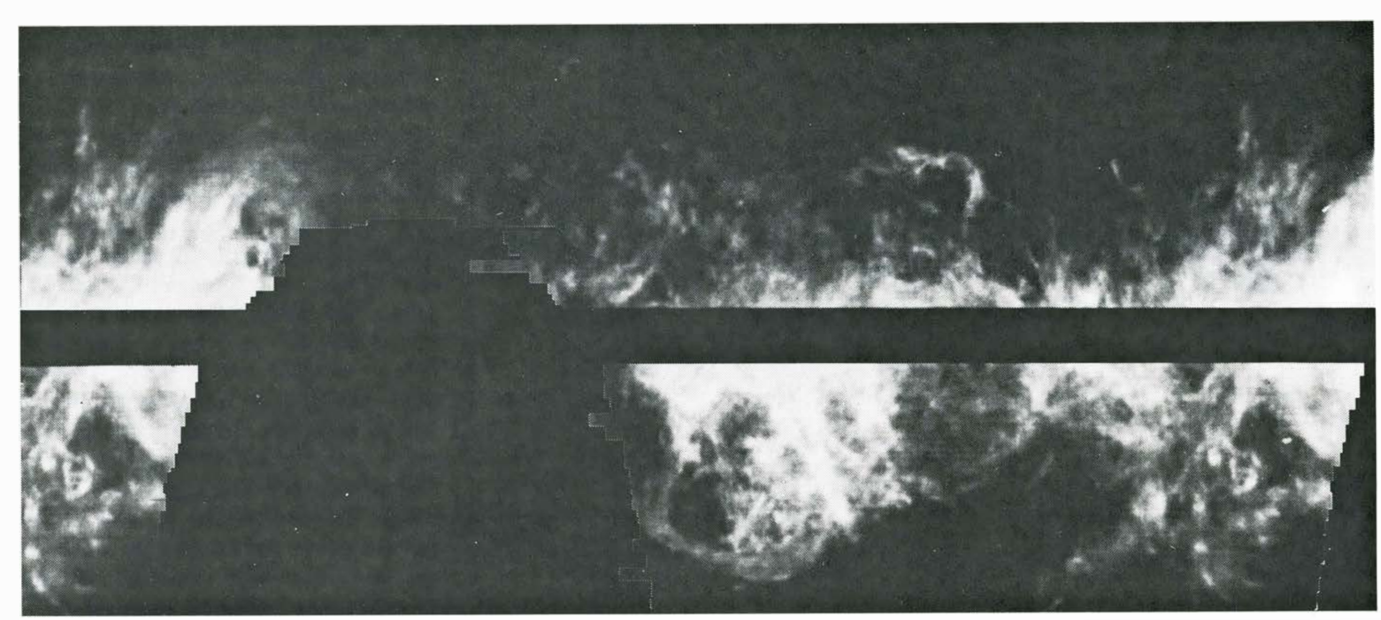
Al sacar nuestra primera fotografía sintética descubrimos casualmente un método menos directo para asociar las observaciones de estrellas con las referentes al medio interestelar. Algunos filamentos del gas aparecen alineados con las líneas de fuerza del campo magnético de la galaxia. El campo magnético alinea los prolongados granos de polvo interestelar, que están asociados con el gas, de suerte que el polvo actúe como un filtro que polarice la luz estelar que pase a través de él. Se cree que la dirección de polarización es paralela a las líneas de fuerza del campo magnético; por consiguiente, una medición de la dirección en que se halla polarizada la luz estelar es de hecho una medición de la dirección del campo magnético galáctico, promediada sobre las distintas concentraciones de material interestelar que se encuentra entre el sistema solar y la estrella emisora de la luz. Si sólo existe una con-

centración de material interestelar entre el sistema solar y la estrella, la dirección en que está polarizada la luz de ésta indica inequívocamente la dirección del campo magnético. La polarización de la luz estelar se ha medido para un gran número de estrellas distribuidas por todo el cielo. Los resultados han sido recopilados y organizados en un mapa por D. S. Mathewson y K. L. Ford, de los Mount Stromlo y Siding Spring Observatories, de la Australian National University *(véase la ilustración de las páginas 58 y 59)*.

A pesar del gran número de estrellas registradas en los mapas de polarización, existen regiones del espacio inadecuadamente cubiertas. Hay también muchas zonas donde la polarización se presenta de forma caótica, debido probablemente a la existencia de más de una concentración de material interestelar en la misma dirección de observación. Estos problemas hacen difícil avanzar afirmaciones generales sobre la relación exacta entre el gas interestelar y el campo magnético galáctico. A pesar de todo, hay una región próxima a las latitudes galácticas positivas e inmediata a la longitud galáctica de 30 grados, donde la alineación del gas con el campo magnético es asombrosa. A partir de esta relación se puede concluir, en mi opinión, que los filamentos del gas están realmente alineados con el campo magnético galáctico.

Las distancias de las estrellas de esta región indican que el gas dista del sistema solar unos 100 parsecs. Los largos filamentos arqueados de gas presentan un tamaño angular de 50 grados, lo que supone que su longitud debe ser del orden de los 100 parsecs. La anchura se extiende sólo hasta unos pocos grados, lo que equivale aproximadamente a unos 14 parsecs, que era el diámetro de una "nube convencional" predicho por el modelo de nubes. Parece altamente improbable que el acuerdo entre la anchura de los filamentos y la predicción del modelo de nubes se deba al azar.





UN ANILLO ELIPTICO DE GAS se puede observar en el centro de la zona inferior de estas dos fotografías sintéticas, reproducidas aquí en negativo. La superior es una fotografía sintética del gas contenido en un estrecho intervalo de velocidades comprendido entre más 16,4 y más 18,4 kilómetros por segundo; la inferior es una fotografía sintética del gas correspondiente al intervalo de velocidades entre más 3,7 y 5,7 kilómetros por segundo. El anillo elíptico es algo mayor en la fotografía inferior, realizada para velocidades inferiores, que en la superior.

Ya que las observaciones visuales del gas interestelar dependen de la visión del gas iluminado por las estrellas de fondo, y dado el escaso muestreo que puede obtenerse del material interestelar con esta técnica, es muy poco probable observar dos o más estrellas que presenten la misma alineación que la dirección del filamento.

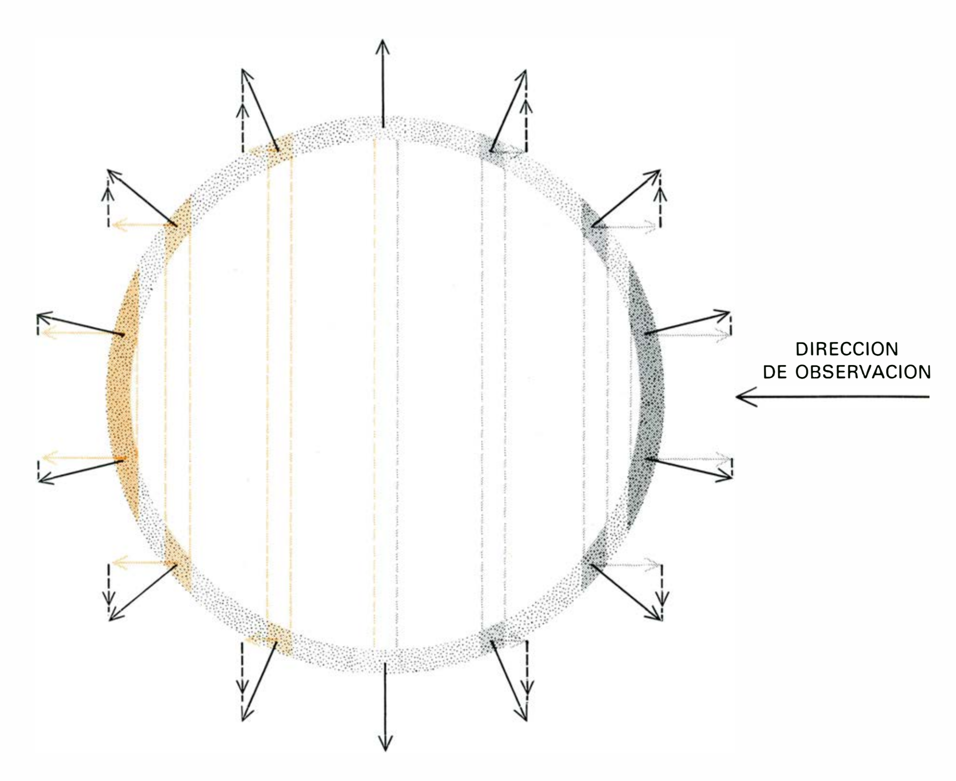
Por tanto, las observaciones visuales del medio interestelar dan un tamaño de las formaciones gaseosas, que equivale, más que a la longitud, a la anchura del filamento.

Mezclas de filamentos

La primera fotografía sintética de la cantidad total del gas del espacio interestelar nos descubrió muchos aspectos de la estructura del medio interestelar. Mis colegas y yo mismo pensamos que sería posible aprender aún más sobre el gas si nos poníamos a examinar la forma como se movían los filamentos. Elegimos diferentes colores para representar velocidades distintas del gas. Siguiendo la convención de que la longitud de onda de la radiación emitida por un objeto que se aleja del sistema solar presenta un corrimiento Doppler hacia la zona roja del espectro, elegimos el rojo para representar el gas que se mueve alejándose del sistema solar, el amarillo-verde para indicar el gas que no muestra corrimiento Doppler y el azul para representar el gas que se acerca al sistema solar. Hicimos una representación del gas presente en un extenso intervalo de velocidades, entre menos 20 kilómetros por segundo y más 20 kilómetros por segundo; la intensidad de cada color representaba la cantidad total del gas que se movía a velocidades comprendidas dentro de dicho intervalo.

El rasgo más sobresaliente, a gran escala, de esta fotografía sintética en color era el siguiente: para latitudes galácticas bajas, el color del gas variaba del rojo hacia el azul al aumentar la longitud galáctica; esta variación de color resulta de la rotación diferencial de la galaxia. La galaxia no tiene un movimiento de rotación semejante al de un sólido rígido, como, por ejemplo, un disco de fonógrafo; de serlo, no existiría movimiento de una parte de la galaxia con respecto a otra.

En cambio, las regiones más internas de la galaxia giran más rápidamente que las externas. Esta rotación diferencial implica que las estrellas más alejadas tienen movimientos grandes con respecto a las que se encuentran más cercanas entre sí. Al comparar las observaciones de las velocidades diferenciales de estrellas con respecto al sol, a gran escala, con la velocidad media del gas interestelar, podemos efectuar una estimación aproximada de la distancia del gas al sistema solar. La estimación de la distancia con este método concuerda con las realizadas independientemente por otros métodos que ya he descrito.

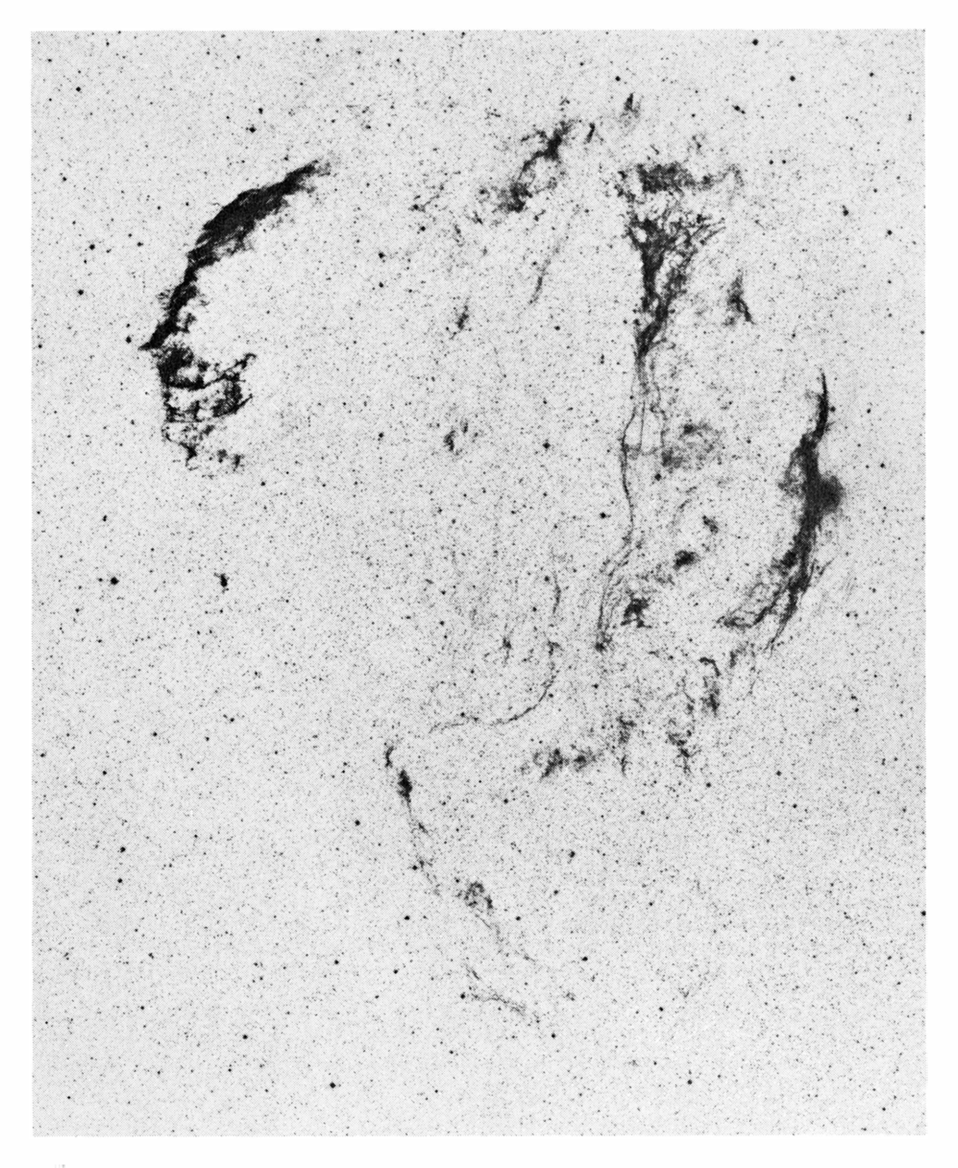


CAPA DE GAS EXPANDIENDOSE RADIALMENTE, en el espacio interestelar; dará lugar a formaciones elípticas o circulares semejantes a las que aparecen en las fotografías sintéticas de la página anterior. En esta imagen presentamos la sección transversal de una capa esférica de gas que se está expandiendo radialmente en todas las direcciones *(flechas negras)*. **El desplazamiento Doppler, mediante el cual mide el observador la velocidad de la capa, depende sólo de la componente radial de la velocidad de la capa** *(flechas grises y en color)*. **Cuando un observador lejano examina una de estas capas en un estrecho intervalo de velocidades, verá un anillo cuyo tamaño depende de la velocidad radial considerada. La velocidad medida será máxima para el punto más próximo** *(gris obscuro)* **y el más alejado** *(color más obscuro)* **de la capa, que están acercándose o alejándose del observador casi directamente; con estas velocidades elevadas, el diámetro de la capa medido será pequeño. Las mínimas velocidades medidas se obtendrán para los sectores próximos al limbo de la capa** *(gris más claro y color menos intenso)*, **que se desplazan perpendicularmente a la dirección de observación; para estas velocidades el diámetro de la capa será grande.**

Lo que desvelaban los distintos filamentos

A una menor escala, la fotografía sintética en color reveló que el gas tiene distribución de velocidades muy detallada. En general, un filamento singular no se caracteriza por una velocidad única. De hecho, nos pareció que no hay una pauta sistemática sobre la forma en que varía la velocidad según la longitud y el grosor del filamento. Unas partes del filamento aparecían formando rizos o lazos, de modo que tales estructuras se caracterizaban, a primera vista, por variaciones al azar en la velocidad. Estas variaciones proporcionaban mucho más información sobre los movimientos en el filamento de la que se disponía hasta entonces, pero era una información molesta para quien pretendiera comprender la estructura global del gas interestelar. Es difícil llegar a obtener una interpretación o explicación correcta de una mezcla de números aleatorios.

En un esfuerzo por dar sentido a la compleja estructura del medio interestelar, decidimos examinar el gas para intervalos de velocidad menores. Dividimos el rango de velocidades de menos 20 kilómetros por segundo a más 20 kilómetros por segundo en 20 intervalos de velocidad de dos kilómetros por segundo; para cada uno, sintetizamos una fotografía del gas. Dados los resultados de la representación en color sobre el intervalo completo que presentaba variaciones aleatorias aparentes en la velocidad sobre tamaños angulares relativamente pequeños, no esperábamos que las series de fotografías para los pequeños intervalos de velocidad dieran resultados satisfactorios. Esta previsión andaba equivocada. Parece que fue este



"LAZO DEL CISNE" (Cygnus Loop), fotografiado para longitudes de onda visibles con el telescopio Schmidt de 1,2 metros del Monte Palomar; es una capa similar a la envoltura elíptica que aparece en la fotografía sintética de la página 60. Como puede verse en este negativo, el Lazo del Cisne no es perfectamente circular, debido a que el medio interestelar que lo rodea no es homogéneo ni se distribuye regularmente. La capa se está expandiendo a una velocidad del orden de 100 kilómetros por segundo y tiene un radio de unos 20 parsecs. Según se deduce, probablemente fue producido por una supernova, una estrella que explosionó catastróficamente hace unos 45.000 años.

paso la clave para la comprensión de la verdadera naturaleza de la estructura del gas interestelar y del tipo de fenómenos físicos que producen tal estructura: nos permitió observar el gas agrupado en una estructura sin interferencias con otras formaciones gaseosas en la misma dirección de contemplación.

No es posible mostrar aquí las veinte fotografías sintéticas, por eso he escogido sólo las dos que, por reunir unas características interesantes, mejor ilustran mi argumentación. Las dos fotografías sintéticas aparecen en la página 60; la superior abarca el intervalo comprendido entre más 16,4 y más 18,4 kilómetros por segundo; la inferior cubre el intervalo entre más 3,7 y más 5,7 kilómetros por segundo. Ambas muestran una gran estructura elíptica de filamentos gaseosos de unos 50 grados de anchura, centrada en las proximidades de la longitud galáctica de 200 grados y la latitud galáctica de –40 grados. De la comparación de las dos fotografías sintéticas se deduce que la estructura con mayor diámetro es la de menor velocidad y la de menor diámetro es la que posee mayor velocidad. Quedamos mudos al localizar una tal estructura coherente en una región del cielo con semejante tamaño angular y tan amplio intervalo de velocidades. Más sorprendente resultó comprobar que esta estructura admitía una interpretación muy sencilla.

Consideremos una capa o envoltura esférica de gas que se está expandiendo radialmente hacia el exterior a partir de su centro. Para un observador exterior a la capa, el corrimiento Doppler del gas que se mueve depende de la proyección del movimiento del gas según la línea visual. El gas que se encuentra en el límite más cercano de la capa, así como el que se halle en el más lejano (según la visual que pasa por el centro de la capa), estarán, respectivamente, acercándose y alejándose directamente del observador y presentarán los mayores corrimientos Doppler. El gas, en el borde de la capa que tiene una mayor distancia angular del centro, estará moviéndose, sin embargo, casi en ángulo recto con respecto a la dirección de observación, y tendrá un menor desplazamiento Doppler. El gas situado entre el centro y el limbo de la capa poseerá velocidades radiales intermedias y le corresponderán desplazamientos Doppler intermedios. Por consiguiente, en cada uno de los intervalos de velocidad más estrechos el observador vería un anillo de gas cuyo tamaño angular decrece conforme aumenta la velocidad radial; tal fue exactamente lo que observamos. Las formaciones filamentosas elípticas que vemos tienen su menor diámetro para la mayor velocidad y su mayor diámetro para velocidades prácticamente nulas. La capa envolvente se expansiona a una velocidad de 23 kilómetros por segundo.

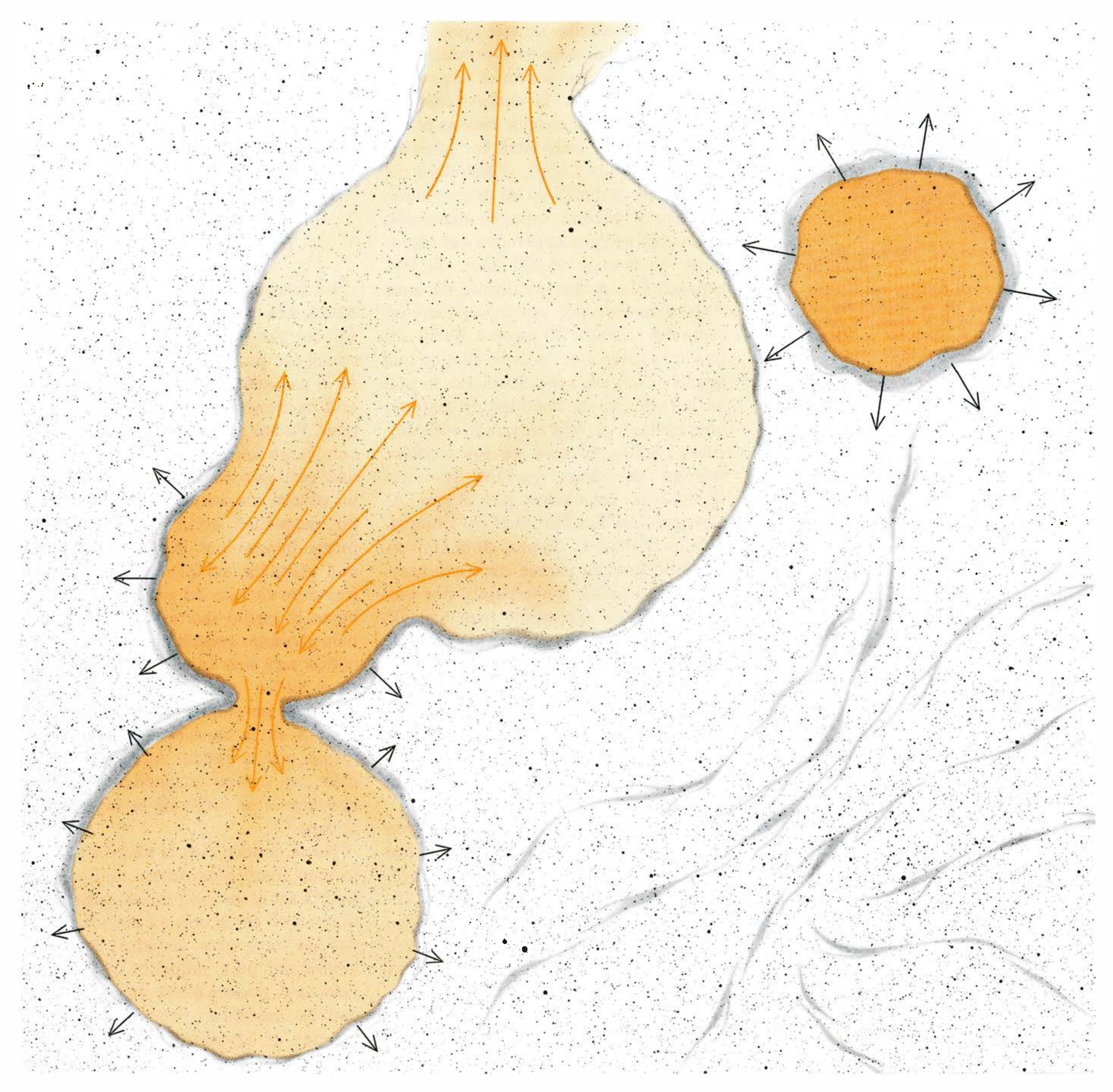
Los filamentos son capas envolventes

En realidad, la capa observada en la fotografía sintética no es una estructura simple e ideal. Para cualquier velocidad el anillo no es circular sino elíptico. No se puede trazar la capa envolvente para el rango completo de velocidades, si bien resulta visible con mayor facilidad a velocidades positivas, que corresponden al ritmo con que se aleja la cara más lejana de la capa. Es preciso recordar, sin embargo, que el medio interestelar no es homogéneo; por consiguiente, no podemos esperar un acuerdo perfecto entre el modelo ideal y la realidad. Las desviaciones respecto del modelo ideal son tan comunes en las observaciones del gas interestelar que, de hecho, nos sentimos afortunados por haber encontrado un buen ejemplo con el que ilustrar el modelo de una capa que se expande radialmente.

Si pudiera determinarse la distancia de la capa elíptica, cabría conocer mejor las propiedades físicas de dicha envoltura. La información en cuanto a la distancia sobre esta estructura concreta es escasa, pero los datos disponibles indican que, salvo un factor 2, dicha distan-

cia es aproximadamente de 150 parsecs. A tal distancia, el tamaño angular de la capa implica un diámetro de 100 parsecs e indica que, a lo largo de su expansión por el espacio, ha barrido una cantidad de materia interestelar igual a 73.000 veces la masa del sol.

Se conocen dos tipos de mecanismos que pueden dar lugar a una gran capa de gas interestelar en expansión. Uno de ellos ha sido sugerido recientemente por John I. Castor, Richard McCray y Robert Weaver, de la Universidad de Colorado; consiste en una rápida transferencia de energía a partir de un viento estelar altamente energético, es decir, un rápido flujo de material que se aleja de la estrella. Se han observado realmente vientos estelares en ciertos tipos de estrellas. Estas estrellas no son muy corrientes, y no existe ninguna estrella



MODELO DE CAPAS de la estructura del medio interestelar, propuesto por el autor para explicar la estructura en filamentos del medio interestelar. Una capa de gas se puede producir bien mediante un viento estelar energético procedente de una estrella caliente, bien mediante una explosión de supernova. Una capa recién formada se expande rápidamente *(arriba, a la derecha)* barriendo el gas del medio interestelar frío que la rodea *(gris claro);* dicho gas es comprimido y calentado hasta formar un denso aro externo de gas ionizado semejante al visible en el Lazo del Cisne *(gris obscuro).* Conforme la capa se expande, deja detrás, en su interior, gas caliente rarificado *(color)* que emite rayos X y produce también líneas de absorción, correspondientes a átomos de oxígeno altamente ionizado, en el espectro ultravioleta de las estrellas calientes. A medida que la capa crece y envejece, el gas se comprime y enfría aún más. Los electrones se recombinan con los iones para formar hidrógeno atómico, el cual emite radiación a la longitud de radio-onda de 21 centímetros. Mientras tanto, se va frenando la expansión de la capa *(abajo, a la izquierda).* Según el trabajo teórico de Donald P. Cox y Barham W. Smith, de la Universidad de Wisconsin, cuando coexisten un gran número de capas en expansión, con el tiempo se producirán colisiones entre ellas. Cuando esto ocurre, los interiores de las envolturas entran en conexión, de suerte que el gas caliente del interior puede pasar de una capa a otra. Si hay suficientes masas en contacto, el gas caliente fluye a través de túneles en el medio interestelar formado por capas densas y frías y por el campo magnético. Una capa muy vieja puede adelgazarse tanto que se rompa por el punto más débil igual que una burbuja *(centro, arriba)* permitiendo que el gas caliente escape al espacio circundante. Los residuos de las capas viejas y deformadas producen jirones de filamentos *(abajo, a la derecha).* (Diseño de A. Beechel).

que presente un viento altamente energético dentro de la capa elíptica de gas observada. La mencionada capa ha podido ser creada por una de estas estrellas que haya completado su evolución y se haya extinguido, pero la energía de la capa es mayor, probablemente, que la que ha podido proporcionar una o incluso varias estrellas de este tipo.

El otro mecanismo para producir una capa de gas en expansión es una explosión de supernova: la desaparición catastrófica de una estrella masiva próxima al fin de su ciclo vital. El "Lazo del Cisne" (Cygnus Loop), o Nebulosa del Velo (Veil Nebula) en la constelación del Cisne, es una capa en expansión producida por una supernova que explosionó hace unos 45.000 años. Se está expandiendo a una velocidad de unos 100 kilómetros por segundo y tiene un radio del orden de 20 parsecs. Es un objeto más joven que la capa elíptica que hemos recientemente descubierto, la cual ha barrido mayor cantidad de material interestelar y se está moviendo más despacio. Sobre la base de los modelos teóricos de capas residuos de viejas supernovas calculados por Roger A. Chevalier, del Kitt Peak National Observatory, y William C. Straka, de la Universidad de Boston, parece claro que nuestra capa elíptica fue expelida por una supernova que explosionó hace 650.000 años y que liberó una energía total de 4×10^{51} ergios. Esta cantidad de energía se halla muy próxima al límite superior de energía liberada en una explosión de supernova generalmente admitido y, probablemente, es varias veces superior a la energía total liberada necesaria para la formación del Lazo del Cisne.

Una capa de material interestelar en expansión, tanto si es creada por viento estelar como si lo es por una supernova, tendrá una estructura característica. Inmediatamente detrás del límite externo de la capa, el gas se comprime y calienta. A medida que el gas calentado es arrastrado por la capa, se mueve lentamente desde el límite externo de ésta hacia el interior de la capa envolvente. Conforme lo hace el gas, se enfría, pero se comprime más todavía. Si la capa resulta de un viento estelar energético, la emisión ultravioleta de la estrella ionizará la región densa que precede (interior) al límite externo.

La capa está constituida por casi toda la materia que originalmente se hallaba contenida en el volumen que encierra ahora la capa. Sin embargo, el interior de la capa envolvente no está totalmente vacío; contiene una pequeña cantidad de gas interestelar original, que ahora se encuentra muy caliente, rarificado y a gran presión. Este gas rarificado es la fuente de la emisión de rayos X que se observa generalmente en los residuos de supernova. El gas rarificado constituye también, a buen seguro, la fuente de oxígeno tan altamente ionizado que ha perdido cinco de sus electrones. Este oxígeno, conocido por O VI, se puede observar a menudo en el espectro ultravioleta de estrellas calientes obtenido con el telescopio del satélite Copérnico. Ambos fenómenos, emisión ultravioleta de O VI y emisión de rayos X, han podido ser comprobados en el interior de nuestra capa elíptica.

Estructura de las capas

La forma en que la materia se concentra en esta capa sugiere que los filamentos del gas interestelar están alineados con el campo magnético galáctico. Incluso el gas más frío del medio interestelar está ligeramente ionizado por la luz estelar. Puesto que los iones son partículas con carga eléctrica se moverán como una corriente eléctrica en presencia de un campo magnético. El gas y el campo magnético están entonces íntimamente ligados por la ligera conductividad eléctrica del primero. Y como el frente de choque del gas que se expande barre el medio interestelar, barre también el campo magnético. El campo magnético será más intenso en la capa que en el medio interestelar circundante, en la misma proporción en que la densidad del gas en la capa es superior a la del medio que la rodea. Además, la dirección del campo magnético tenderá a apoyarse a lo largo de la superficie de la capa.

No se conoce bien la manera como se forman filamentos dentro de una capa. Podría ser el mismo proceso que crea los filamentos en la corona del sol. En todas las fotografías sintéticas del medio interestelar para pequeños intervalos de velocidad pueden verse una gran cantidad de filamentos; hay muchos más de los que se aprecian en las fotografías que he presentado aquí. La mayoría de ellos no presentan cambios de su estructura con la velocidad. No obstante, mis compañeros y yo consideramos los filamentos como porciones de las capas en expansión que se han visto frenadas por colisiones con el gas interestelar ambiental o con otras capas envolventes.

Nuestra interpretación es una extrapolación de las relativamente fiables observaciones de las pocas capas que presentan tamaños diferentes para distintos intervalos de velocidad, pero pensamos

que está justificada. Una de las razones es que en el caso de una capa que muestra un pronunciado cambio en el diámetro con respecto a la velocidad y en el caso de otras dos capas que casi no exhiben cambios del diámetro con la velocidad, hay pruebas independientes, en la forma de la longitud de radio-onda de la emisión de sincrotón, de que las estructuras proceden de una explosión de supernova. La emisión de sincrotón está generada por electrones altamente energéticos que describen trayectorias espirales en el campo magnético interestelar con velocidades cercanas a la de la luz. Este tipo de emisión está asociado muy frecuentemente con capas producidas por supernovas: se cree que en la explosión de supernova original los electrones fueron acelerados a elevadas energías.

La cantidad de filamentos visibles en las fotografías sintéticas implican la existencia de muchas capas en el espacio interestelar. La existencia de tan gran número de capas plantea interesantes ramificaciones que han sido exploradas teóricamente, en primer lugar, por Donald P. Cox y Barham W. Smith, de la Universidad de Wisconsin. Si hay muchas capas en el espacio interestelar es probable que colisionen al expansionarse.

Tras la colisión, las zonas interiores de las capas entran en contacto y el gas caliente del interior de las envolturas puede pasar rápidamente de una a otra. En una situación de equilibrio estacionario, cabe pensar que se habrán unido cierto número de capas formando túneles de gas caliente de baja densidad, rodeados de estrechas capas de gas frío con una alta densidad y de campos magnéticos. Ocasionalmente, una capa puede reventar como una burbuja, si no hay la suficiente materia fría para resistir la presión del gas caliente. La ruptura de la capa proporcionará una "escotilla" de salida para el gas caliente confinado en el interior, que fluirá entonces fuera del espacio limitado por las capas.

El modelo de capas

La exploración de longitud de radioonda que hemos llevado a cabo parece ser coherente con la anterior descripción de la estructura del medio interestelar; las observaciones de rayos X y de las emisiones ultravioletas de O VI del espacio interestelar también parecen confirmarla. El modelo de capas para la estructura del medio interestelar es radicalmente diferente del modelo de nubes y todavía no está totalmente aceptado.

Quedan cuestiones importantes por resolver antes de aceptar esta descripción como enteramente correcta. ¿Concuerdan con este modelo las observaciones realizadas en el espectro visible o proporcionan datos sobre algún otro componente del gas interestelar? ¿Son la estructura de la temperatura y la densidad del gas interestelar observado coherentes con lo que se debe esperar detrás del límite externo de la capa? ¿Concuerdan con dicha interpretación la topología e intensidad de los campos magnéticos en el espacio interestelar?

El nuevo instrumental recién construido o en fase de construcción desempeñará un papel crucial en los esfuerzos necesarios para resolver estas cuestiones. En particular, el radiotelescopio de 304,8 metros de Arecibo, en Puerto Rico, que ha experimentado recientemente una nueva puesta a punto y el complejo de radiotelescopios de "muy larga base" (Very Large Array) de New Mexico proporcionarán la mayoría de los datos necesarios en longitudes de radio-onda. La información en las longitudes de onda correspondientes a la radiación X y ultravioleta es necesaria para determinar si el gas caliente se halla asociado preferentemente con el interior de las capas. Tales datos se pueden obtener exclusivamente a partir de instrumentos instalados en el espacio exterior. Mientras tanto, en la superficie terrestre, los astrónomos están observando el medio interestelar en el espectro visible a través de nuevas técnicas que proporcionan datos de mucho más alta calidad, que pueden ser mejor y más rápidamente comparados con los obtenidos en la zona de longitudes de radioonda.

Quedan pendientes muchos interrogantes a propósito del origen de las capas. La energía que se requiere para crearlas es grande, mayor posiblemente que la que pueden proporcionar vientos estelares o explosiones de supernovas. Al parecer el número de tales capas es elevado. Aunque nadie ha analizado aún esta multiplicidad detalladamente, es muy posible que existan demasiadas envolturas en comparación con los ritmos generalmente aceptados de formación de estrellas con vientos estelares energéticos o con los ritmos a que se producen las explosiones de supernovas. Si esto es así, las estrellas con vientos estelares energéticos o las supernovas son quizá más frecuentes de lo que se supone. O tal vez existe otra clase de objetos, desconocidos hasta ahora, que inyectan grandes cantidades de energía en el medio interestelar.

El mecanismo de adherencia en las bacterias

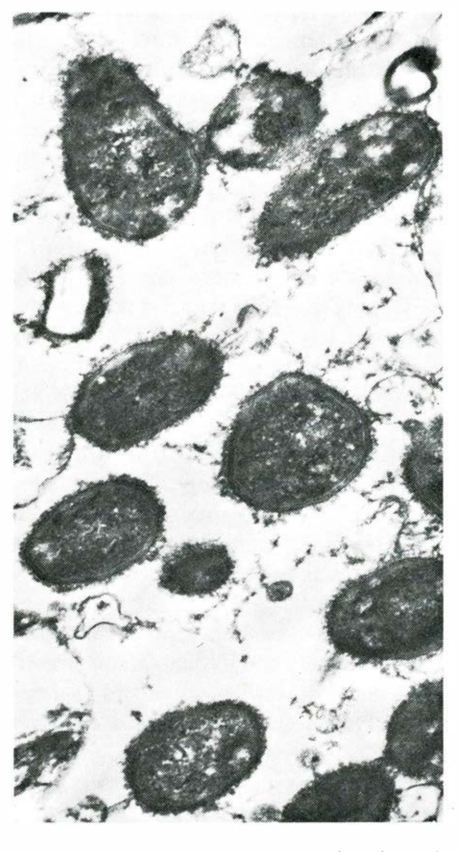
En la naturaleza, pero no en los cultivos de laboratorio, las bacterias están rodeadas por un "glicocálix" de fibras que permite su adherencia a las superficies inertes y a otras células. Esta adherencia puede evitarse con una nueva clase de antibiótico

J. W. Costerton, G. G. Geesey y K.-J. Cheng

Las bacterias pueden fijarse tenazmente, y a menudo con exquisita especificidad, a distintas superficies, desde las piezas dentales y el aparato respiratorio del hombre y el intestino de un bóvido hasta las rocas sumergidas en corrientes de aguas muy veloces.

Esta adherencia se consigue mediante una masa de fibras enmarañadas, constituidas por polisacáridos de molécula ramificada que se extienden desde la superficie bacteriana y forman un compacto glicocálix que rodea a las células individualmente o a una colonia de ellas.

La adherencia proporcionada por el glicocálix da lugar a hábitats especializados de bacterias en muchos ecosistemas naturales; en concreto, el glicocálix es un determinante crítico para la iniciación y progresión de varias enfermedades producidas por las bacterias, como la caries dental y la neumonía.



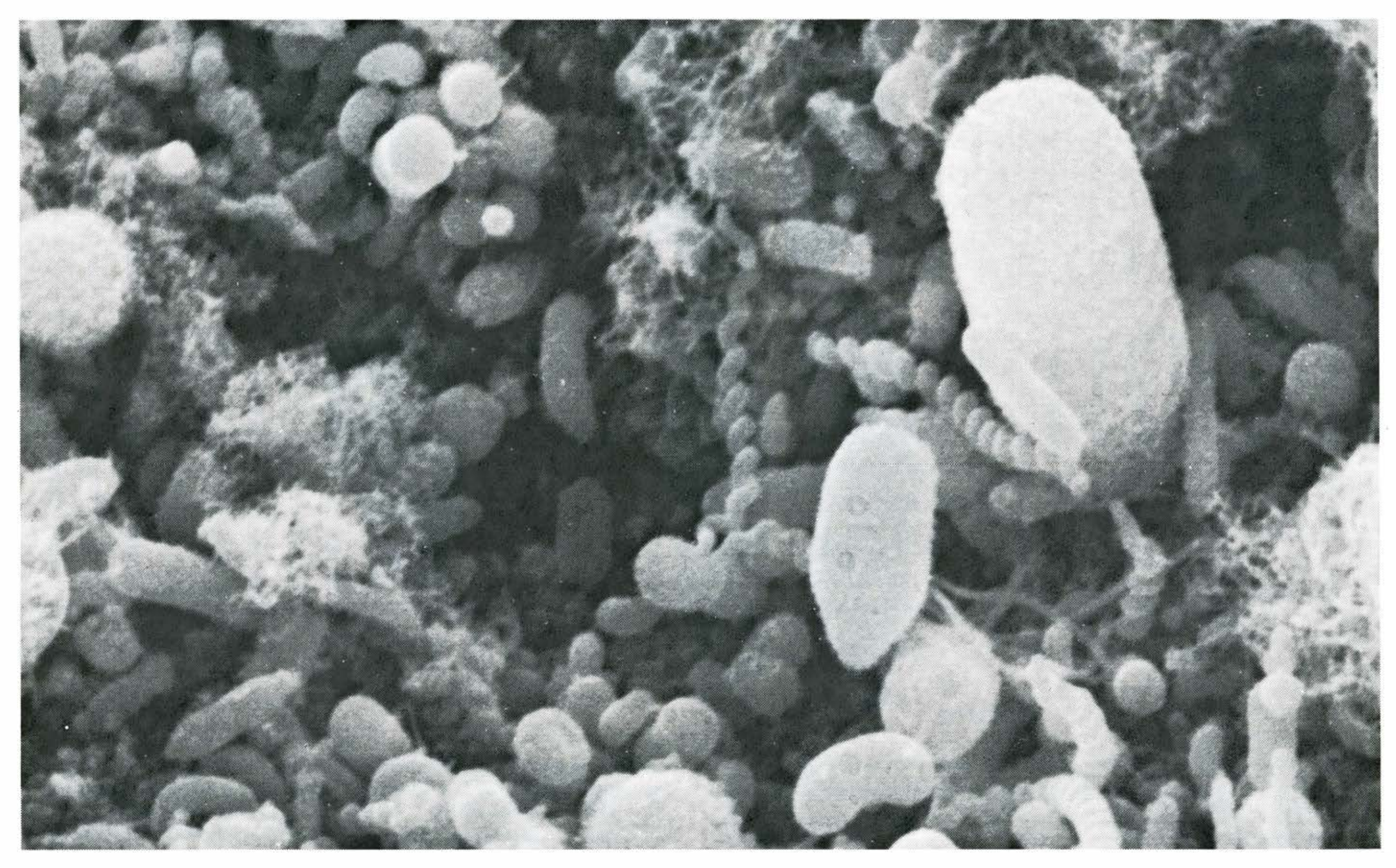
LAS BACTERIAS "DESNUDAS" *(izquierda)* **proceden de un cultivo puro de** *Escherichia coli* **realizado en laboratorio; las bacterias cubiertas por un glicocálix** *(derecha)* **son células de** *Pseudomonas* **obtenidas de una vejiga urinaria humana infectada. En ambos casos, las células se han teñido con rojo de rutenio, tinción específica para las fibras de polisacárido. El glicocálix bacteriano ha sido ignorado hasta hace poco tiempo debido a que las cepas bacterianas de laboratorio que nos son familiares no forman el glicocálix, al no ser necesario en las condiciones habituales de cultivo puro.**

A tan importantes y, tomadas las cosas en su perspectiva, obvios hallazgos en la superficie de la célula bacteriana no se llegó hasta la última década. Paradójicamente, la razón principal de este descubrimiento tardío del glicocálix y sus funciones se ha debido a la total confianza que los microbiólogos han depositado en un sistema experimental enormemente práctico y eficaz: el cultivo puro de una cepa bacteriana en los medios convencionales de laboratorio. Para generar y sostener un glicocálix, la célula bacteriana debe gastar energía; en el medio protegido de un cultivo puro, el glicocálix supone un derroche metabólicamente caro, que no confiere ventaja selectiva alguna. Las células que fabrican estas complejas envolturas suelen eliminarse de los cultivos puros, sustituyéndose por mutantes carentes de glicocálix, que pueden destinar más energía de sus reservas para la proliferación. Los microbiólogos han estudiado durante mucho tiempo esta clase de mutantes "desnudos".

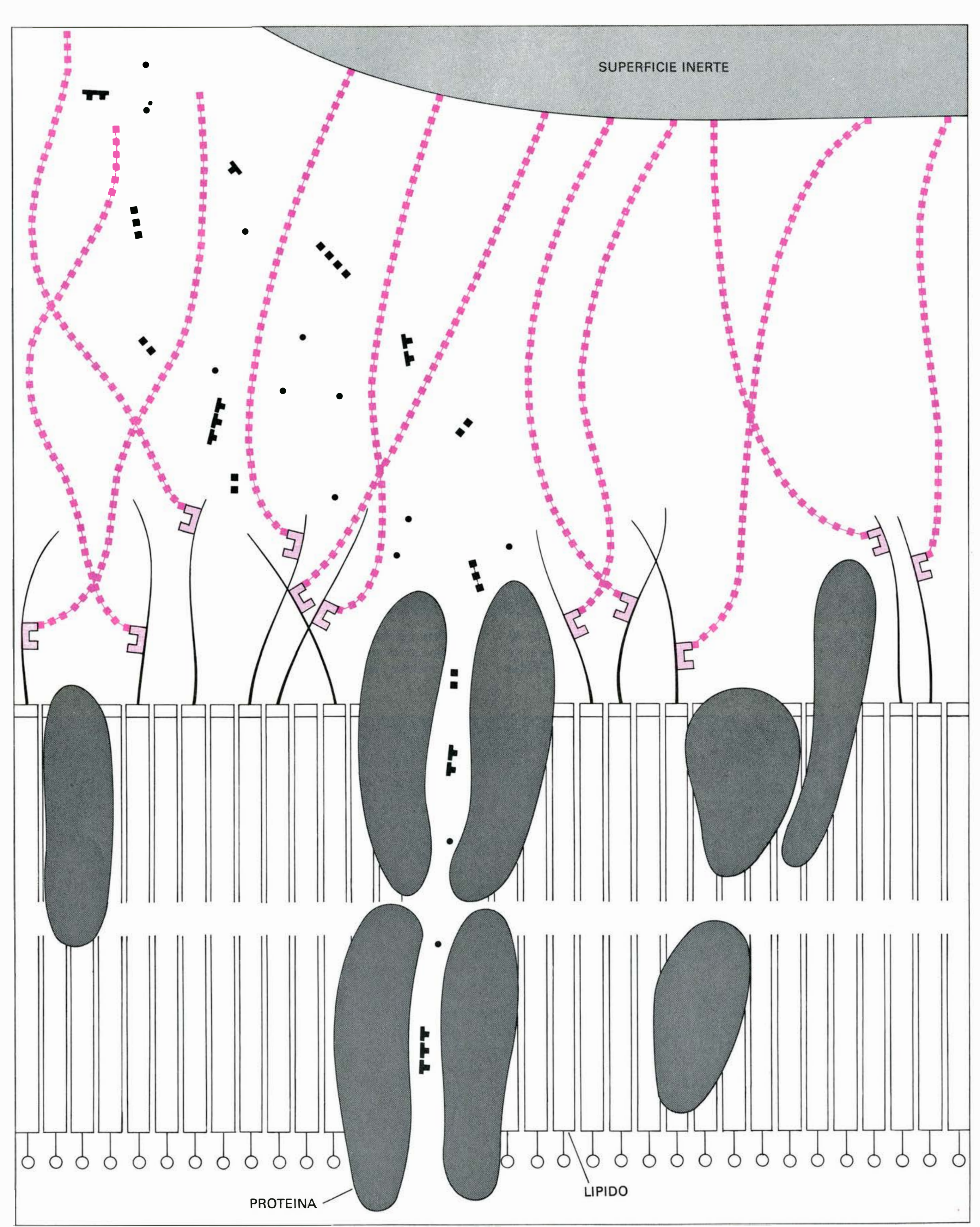
Contrariamente, en ambientes naturales competitivos, poblados por varias especies bacterianas, la selección favorece a las células que están protegidas y capacitadas para adherirse a las superficies adecuadas mediante un glicocálix. En 1969, Ivan L. Roth, de la Universidad de Georgia, demostró la existencia de las fibras de carbohidrato que rodean a las bacterias que habitan en ambientes acuáticos.



EL GLICOCALIX BACTERIANO consiste en un entramado de fibras que se extiende a partir de la superficie bacteriana. Las fibras se adhieren entre sí, a la superficie de otras células e incluso a superficies inertes. La micrografía electrónica, realizada en el laboratorio de Costerton, muestra una población de bacterias fijadas entre sí y a la superficie celular de la panza de un bóvido, a través de sus glicocálices.



POBLACION MIXTA, constituida por diferentes especies bacterianas y algunas células de levadura adheridas a la mucosa de la panza de un bóvido. En esta preparación no se observan fibras de polisacárido aisladas que fijen a los microorganismos, pero algunas fibras del glicocálix están coaguladas y pueden verse como masas de un material espumoso y ramificado. Fotografía obtenida con microscopio electrónico de barrido.

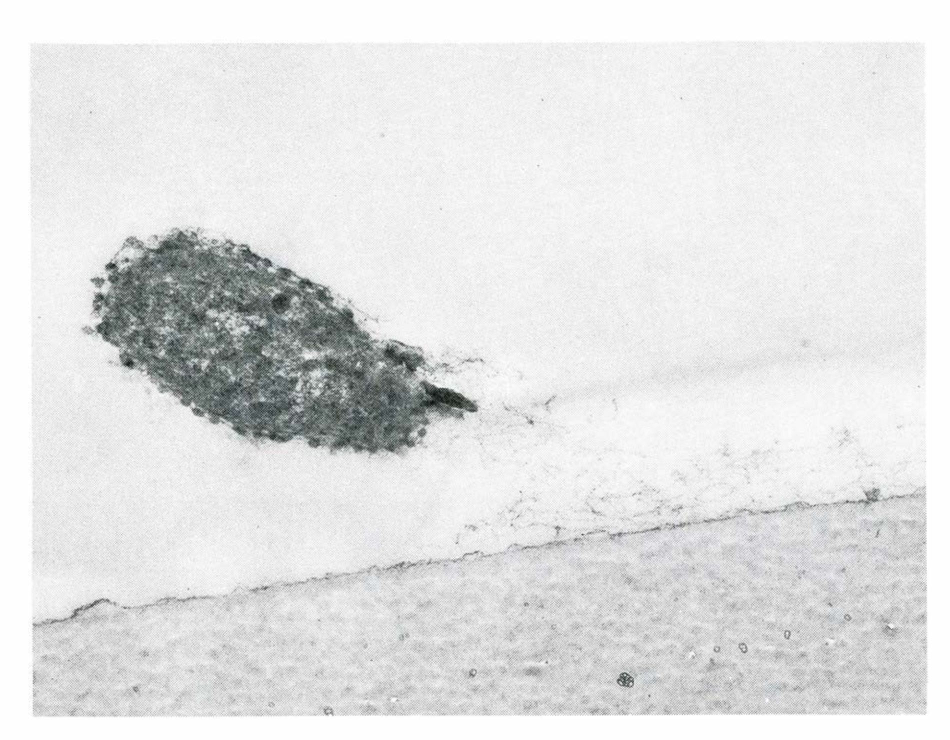


EL GLICOCALIX se extiende a partir de la envoltura celular externa de la bacteria. La membrana celular está constituida por una doble capa de moléculas lipídicas *(estructuras en forma de horquilla)*, entre las cuales se sitúan moléculas proteicas *(manchas en gris)*. Las moléculas de lipopolisacáridos *(filamentos de color negro)* se extienden a partir de la membrana. El glicocálix está formado por una masa de largas fibras de polisacárido *(cadenas de cuadrados coloreados)*. Las fibras están constituidas por cadenas de monosacáridos sintetizadas por los enzimas bacterianos denominados polimerasas *(estructuras en forma de* C*)* asociadas a los lipopolisacáridos. Las fibras del glicocálix se adhieren a las superficies próximas a las células, en este caso una superficie inerte *(arriba, a la derecha)*. Además, las fibras del glicocálix canalizan varias clases de nutrientes hacia la célula bacteriana. Estos nutrientes pueden ser: azúcares *(rectángulos)*, aminoácidos *(figuras en forma de* T*)* e iones inorgánicos *(círculos)*. Estos nutrientes penetran en el interior de la célula a través de conductos de la membrana celular, flanqueados por moléculas proteicas.

En el mismo año, Ian W. Sutherland, de la Facultad de Medicina de la Universidad de Edimburgo, caracterizó los polisacáridos de superficie de bacterias recién extraídas de ambientes naturales, señalando de este modo la universalidad de lo que conocemos por glicocálix. Los estudios realizados en nuestros laboratorios de la Universidad de Calgary y de la Estación Agrícola Experimental Canadiense de Lethbridge, en Alberta, así como en otros centros, han puesto de manifiesto a nivel crítico que el glicocálix es esencial para el éxito biológico de muchas de las bacterias que se observan en la mayoría de hábitats naturales.

Estas superficies recubiertas por polisacáridos no son exclusivas de la célula bacteriana. En realidad, la pared celular de los vegetales superiores constituida por polisacáridos más compactos fue una de las primeras estructuras microscópicas descritas por Robert Hooke en 1665. La superficie análoga en las células animales, muy semejante al glicocálix bacteriano, fue descrita en 1971 por Vincent T. Marchesi y sus colaboradores del Instituto Nacional de Artritis y Enfermedades Metabólicas y Digestivas. Estos autores aislaron e identificaron las glicoproteínas de la membrana de las células animales y mostraron que las fibras de polisacárido que forman parte de dichas glicoproteínas se extienden desde la membrana celular hacia fuera; estas proyecciones se fusionan, constituyendo un glicocálix que representa, de hecho, la verdadera superficie activa de la célula. La naturaleza exacta de este ubicuo glicocálix varía según la composición de monosacáridos de las fibras, de modo que difiere de una célula a otra de la anatomía del organismo y, probablemente, varía asimismo durante los distintos estadios del ciclo vital de cada célula.

La sorprendente conclusión que se extrae de estas observaciones es que la superficie activa real de cualquier célula bacteriana, vegetal superior o animal, consiste en un entramado compacto de fibras de polisacárido construidas y orientadas por la propia célula. Una célula bacteriana se adhiere a una célula animal, una célula vegetal o incluso a otra bacteriana, por yuxtaposición de su propio glicocálix a la superficie de la célula a la que se adhiere. En muchos casos, el nexo de unión es suministrado únicamente por unas proteínas denominadas lectinas, que se unen específicamente a polisacáridos de una estructura

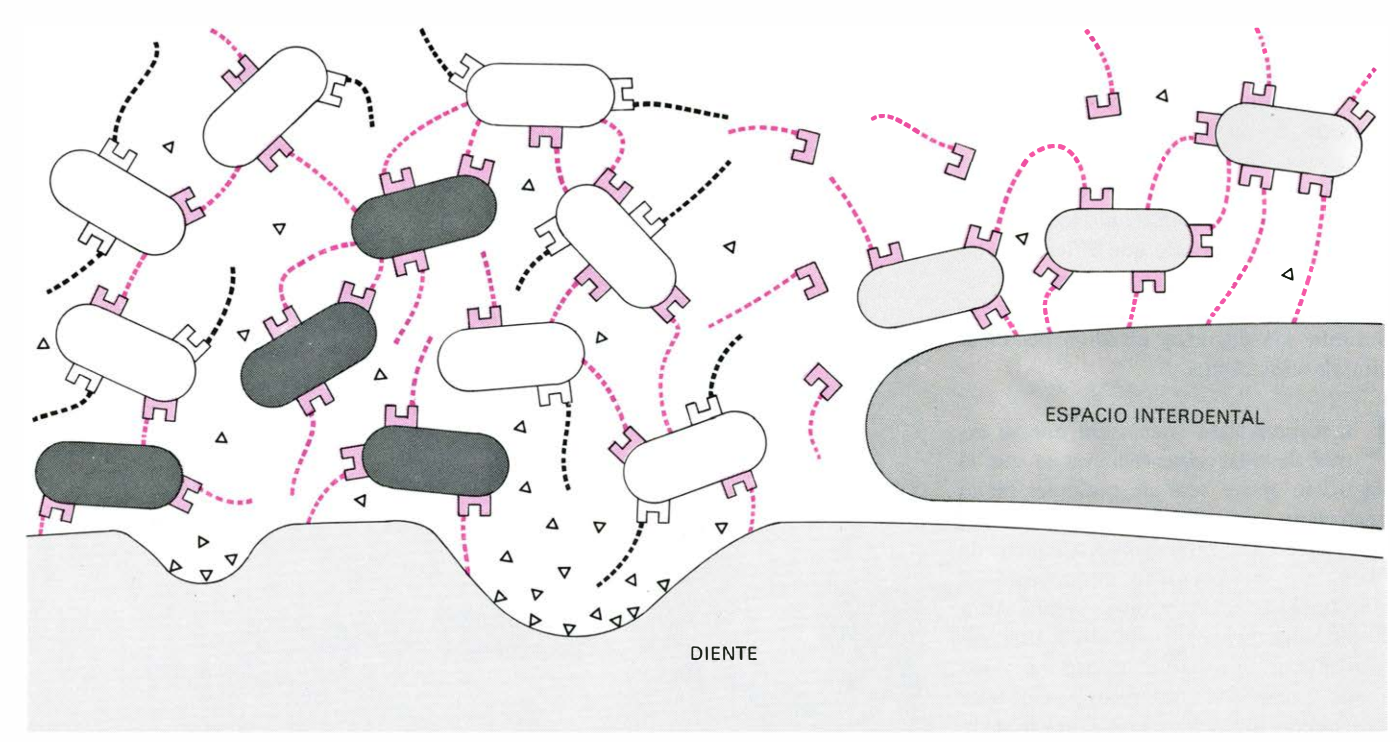


CELULA BACTERIANA sujeta, por medio de las fibras de su glicocálix, a la superficie de un disco de plástico sumergido en agua corriente de un torrente de montaña, salvándose del arrastre.



GLICOCALIX ELABORADO, que permite la fijación de una célula bacteriana *(arriba)* al glicocálix de una célula animal. Este aparece por encima de la membrana plasmática de doble capa.

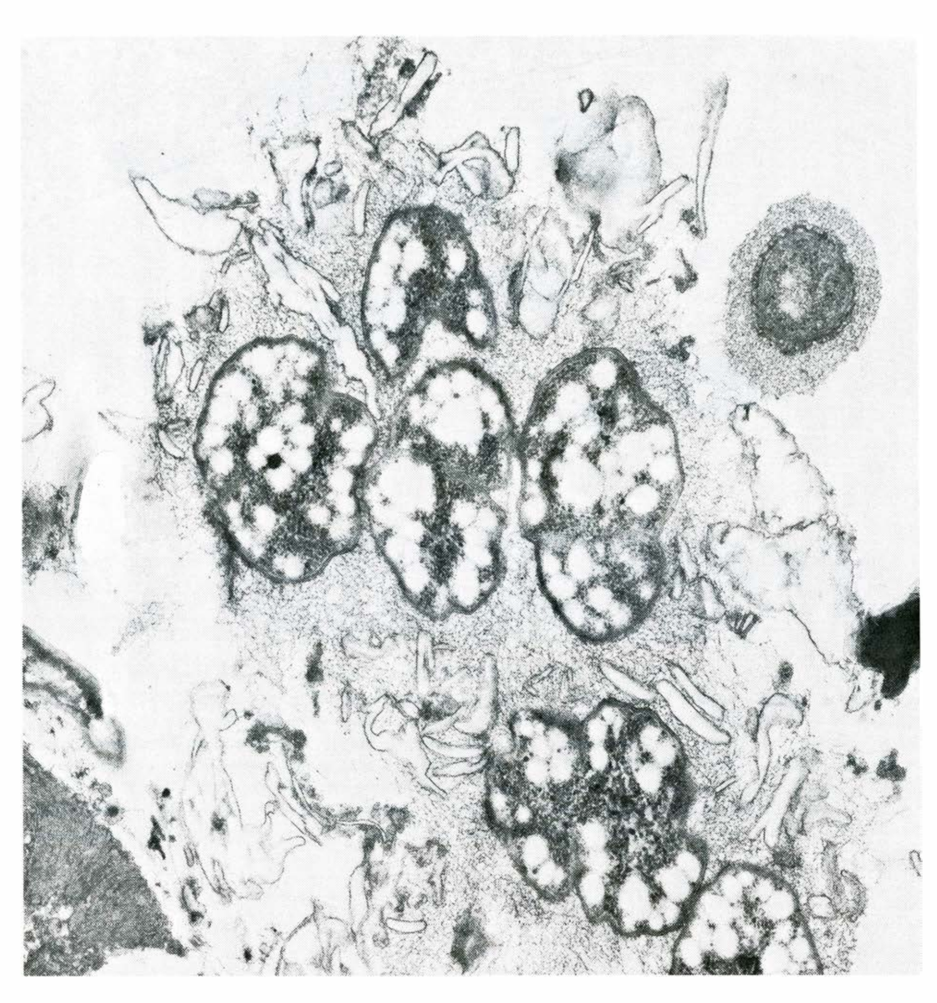
EL PATOGENO DENTAL *Streptococcus mutans* **interactúa con la sacarosa o azúcar común del modo que se muestra en este esquema, realizado según las observaciones de Ronald J. Gibbons, del Centro Odontológico Forsyth en Boston. La sacarosa está compuesta de una molécula de glucosa** *(cuadrados coloreados)* **y otra de fructosa** *(cuadrados negros).* **El enzima invertasa** *(en el centro)* **disocia la sacarosa, liberando sus dos componentes que penetran en la célula para ser utilizados como nutrientes. La glucosiltransferasa** *(izquierda)* **disocia la sacarosa liberando las moléculas de fructosa y uniendo las unidades de glucosa en un polímero de gran longitud denominado glucano, insoluble en agua, y que constituye el principal componente del glicocálix. Por su parte, el enzima fructosiltransferasa** *(derecha)* **libera la glucosa y polimeriza la fructosa.**



EN LAS PIEZAS DENTALES se forma una película amarillenta, denominada sarro dental, que está compuesta por bacterias fijadas en un entramado de material del glicocálix. Algunas bacterias de la cavidad bucal *(izquierda, de color gris obscuro)* **fabrican sus propios glucanos adherentes. Otras bacterias como** *Streptococcus salivarius (de color gris claro)* **colonizan los espacios interdentales** *(derecha),* **fabricando glucanos que se liberan y desplazan hacia las piezas dentales próximas, llevando asociados enzimas polimerizantes. Una vez en la superficie dental, estos glucanos libres pueden contribuir a la formación de un entramado que engloba a varias bacterias no específicas** *(de color blanco),* **originando el sarro. En el interior de éste, los enzimas bacterianos** *(triángulos)* **capaces de erosionar el esmalte se dirigen hacia el diente causándole caries.**

molecular muy particular [véase "Lectinas", por Nathan Sharon; INVESTIGACION Y CIENCIA, agosto, 1977].

Los primeros estudios sobre la formación del glicocálix en las bacterias se llevaron a cabo con *Streptococcus mutans,* un microorganismo que coloniza los dientes de los seres humanos, y fueron realizados por Ronald J. Gibbons, del Centro Odontológico Forsyth en Boston, y otros investigadores. Entre 1960 y 1967, estos autores describieron que había tres enzimas localizados en la superficie de la bacteria que interactuaban muy particularmente con el azúcar común, la sacarosa, compuesta por una molécula de glucosa y otra de fructosa. La invertasa es un enzima que hidroliza la sacarosa separándola en sus dos componentes que, al ser liberados, constituyen una fuente de energía para la célula bacteriana. Un segundo enzima, la glucosiltransferasa, rompe la molécula de sacarosa y libera la fructosa como nutriente, pero polimeriza la glucosa en un polisacárido de cadena larga denominado glucano, insoluble en agua. El tercer enzima, la fructosiltransferasa, polimeriza la fructosa convirtiéndola en un polisacárido soluble en agua y libera, en cambio, la glucosa.

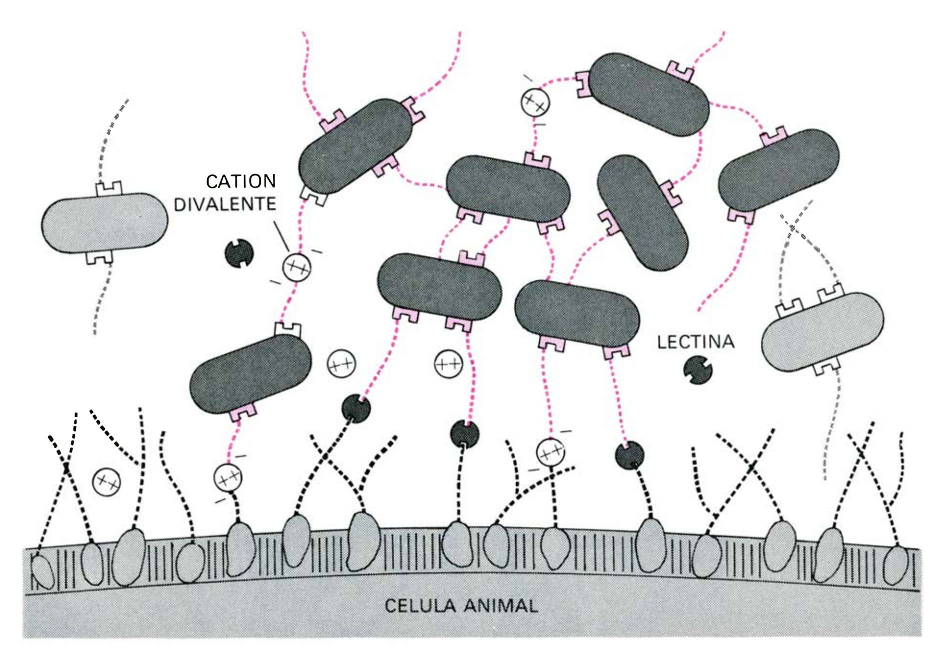


LAS DIVISIONES CELULARES BACTERIANAS en el interior del glicocálix tienden a producir microcolonias, constituidas por células de la misma especie. Esta micrografía muestra una microcolonia rodeada por un glicocálix que engloba asimismo restos de células muertas, cuyos constituyentes se han utilizado como nutrientes por las células supervivientes. Puede observarse también *(arriba, a la derecha)* **una célula, que no pertenece a la colonia, con su propio glicocálix.**

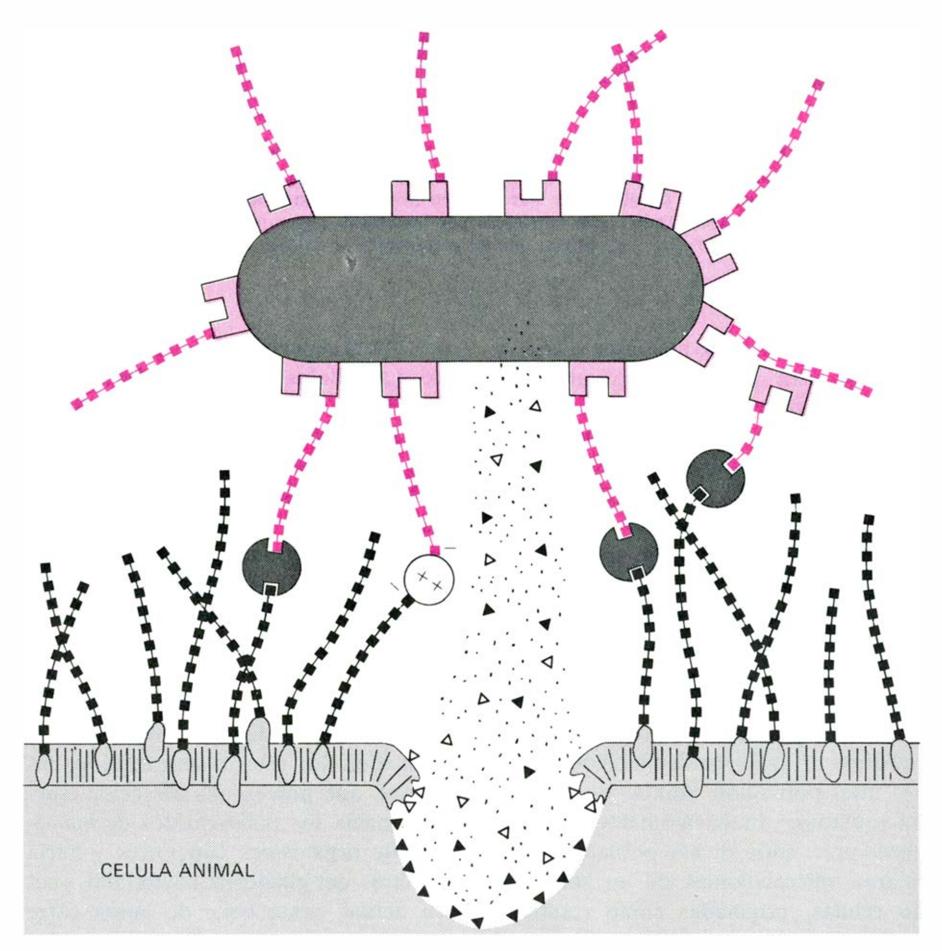
El hecho importante es que, de algún modo, el glucano puede adherirse a la superficie del esmalte dental fijando así la célula bacteriana en la pieza dental. El glicocálix aumenta su grosor continuamente debido a que en el armazón de glucano se halla el enzima glucosiltransferasa. En su crecimiento, el glicocálix engloba a más células bacterianas de la misma o de distinta especie, originando la película amarillenta conocida como sarro dental. Otra bacteria de la cavidad bucal humana, *Streptococcus salivarius,* coloniza los espacios interdentales en lugar de los propios dientes. Esta bacteria libera glucanos que se desplazan hacia las superficies dentales adyacentes y, con ayuda de la glucosiltransferasa que contienen, forman un entramado de polisacárido que aglomera una población mixta de bacterias y, de este modo, originan el sarro dental. En esta estructura se concentran los enzimas que, actuando sobre el esmalte dental, son responsables de la caries.

El estudio de Gibbons acerca de los patógenos dentales dirigió la atención de los investigadores hacia otros hábitats bacterianos. Roth y otros investigadores observaron que una amplia variedad de superficies sumergidas en agua estaban cubiertas por una película de polisacáridos de origen bacteriano. Nosotros investigamos el proceso mediante el cual se formaba esta pátina de polisacáridos, colocando discos de plástico en un torrente de montaña. Al examinar los discos con el microscopio electrónico, hallamos que una bacteria podía fijarse a la superficie inerte del disco de plástico por medio de un haz de fibras de polisacárido trenzadas, que son capaces de resistir enormes fuerzas de cizalla. La colonización inicial de este tipo de superficies se lleva a cabo por bacterias o por algas (que poseen también este tipo de fibras de polisacárido), y con el tiempo se forma un armazón de fibras que incluye una población mixta de células microbianas. Frecuentemente, pueden observarse, entre dichas poblaciones celulares, microcolonias de un solo tipo de células, originadas como resultado de divisiones celulares acontecidas en el interior del glicocálix común.

Se desconoce por el momento cuál es el mecanismo de adherencia que permite la unión de las fibras de glicocálix a una roca o a la superficie de un disco de plástico. Las células de los vegetales superiores o de los animales, al contrario de las superficies inertes antes mencionadas, ofrecen a las bacterias sus paredes celulares o sus glicocálices que están químicamente definidos por la composición de sus propios polisacáridos. Las fibras de polisacárido del glicocálix bacteriano, que se encuentran cargadas negativamente en la mayor parte de su extensión, pueden establecer enlaces iónicos con los polisacáridos de las células eucariotas por mediación de los cationes divalentes del medio. En este mismo sentido, las lectinas, que poseen una atracción específica hacia los polisacáridos de las células de organismos superiores y hacia las fibras del glicocálix bacteriano, pueden actuar como nexo de unión entre estas dos estructuras. Obviamente, las bacterias cuyas fibras no son capaces de unirse ni a los polisacáridos de una célula eucariota ni a las lectinas presentes en el medio, carecen de capacidad



EL GLICOCALIX DE LAS CELULAS ANIMALES está constituido asimismo por una trama de fibras de polisacáridos ramificados, unidas a las moléculas de glicoproteína de la membrana celular *(abajo).* El glicocálix de las bacterias y el de las células animales interactúan entre sí, tal como señala este esquema. Los monosacáridos terminales de las fibras interactúan directamente por medio de enlaces polares, reaccionando los azúcares cargados negativamente con cationes divalentes como el magnesio. El nexo de unión puede establecerse también mediante una lectina *(gris obscuro),* que es una proteína que posee receptores específicos para las fibras de polisacáridos. Las bacterias *(gris claro)* que posean polisacáridos de distinta naturaleza química no se adherirán.

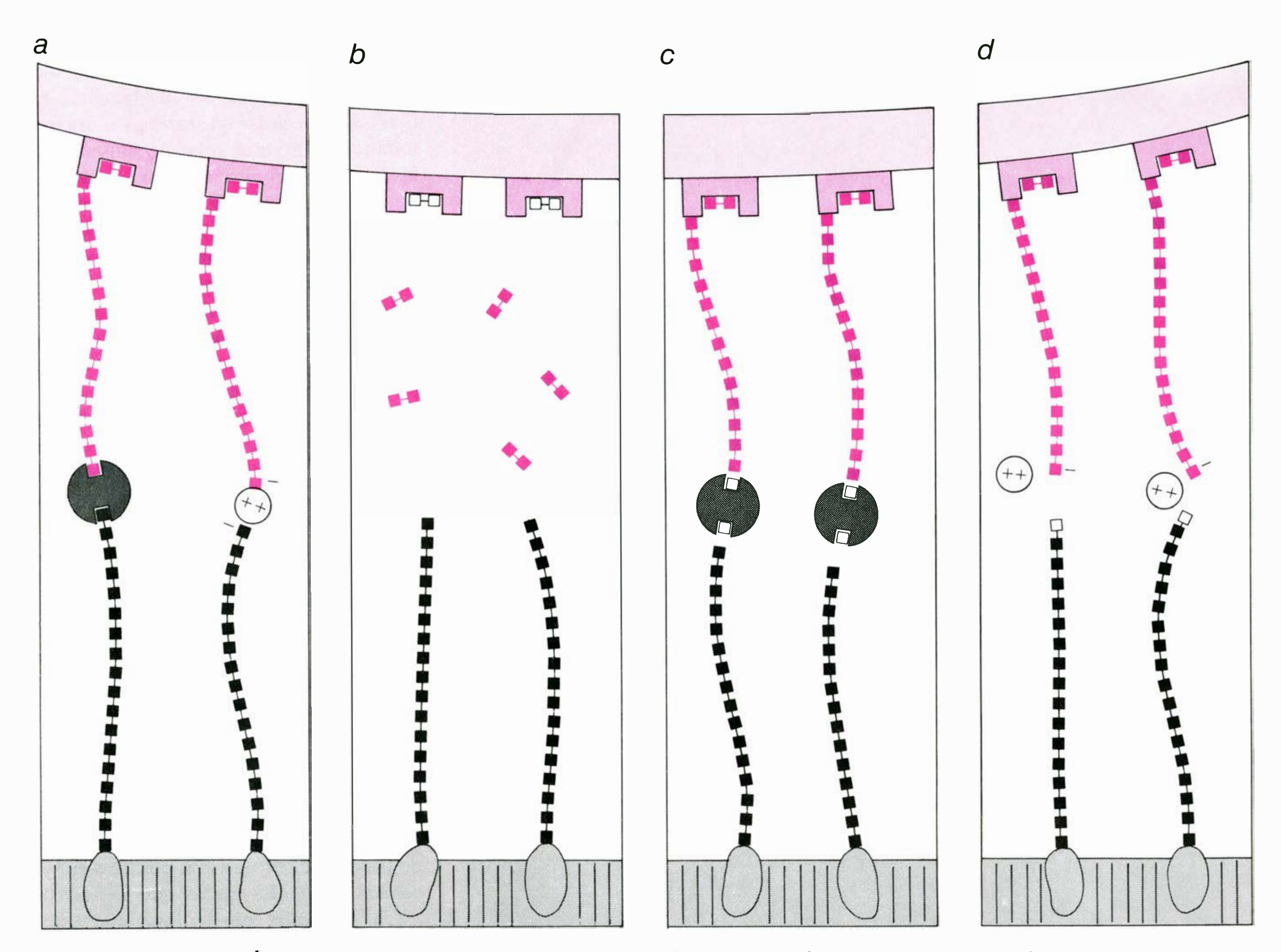


LA INFECCION BACTERIANA puede desencadenarse mediante la adherencia específica a ciertas células del tejido animal. El compacto entramado constituido por las fibras de los dos glicocálices suministra un entorno en el cual los enzimas y toxinas bacterianos *(triángulos)* difunden óptimamente, lesionando la membrana celular eucariota y digiriendo su contenido. Las moléculas nutrientes liberadas de la célula animal *(puntos negros)* se canalizan hacia la bacteria infecciosa.

adherente. Como resultado de la especificidad de la adherencia bacteriana, la superficie celular de las células eucariotas es colonizada frecuentemente por una sola especie de bacterias, que pueden proliferar y formar una microcolonia englobada y fijada por las fibras de las células bacterianas que la forman. Una vez constituidas estas microcolonias, puede ocurrir que bacterias de otra especie, incapaces de adherirse a las células eucariotas, se fijen al glicocálix de los colonizadores iniciales, originando así una población mixta.

La capacidad de adherencia a un tejido orgánico particular de una especie bacteriana determinada puede averiguarse colocando a la bacteria problema en un medio adecuado, dentro de una cámara en la que se ha depositado el tejido en cuestión como base de la misma. Al cabo de cierto tiempo, se examina el tejido para determinar el número de bacterias que se han adherido al glicocálix de las células del mismo. Andrew B. Onderdonk y sus colegas, del Veterans Administration Hospital, en Boston, demostraron que, en condiciones análogas a las descritas, las células de una cepa de *Bacteroides fragilis,* con glicocálix bien desarrollado, se adhieren muy bien al peritoneo de rata, mientras que células mutantes de la misma especie bacteriana, que no producían el glicocálix, no eran capaces de adherirse a dicho tejido. Otras experiencias similares sugieren que la naturaleza del glicocálix de las células eucariotas cambia con la edad celular; este hecho puede explicar por qué ciertas células del intestino de los bóvidos puede ser colonizado con profusión, en tanto que otras células del mismo tejido no pueden serlo. El reciente descubrimiento de que las células infectadas por un virus tengan alterado el glicocálix puede explicar el aumento de susceptibilidad a las infecciones bacterianas que se observa frecuentemente después de una infección vírica.

Si las bacterias nativas de las corrientes rápidas de los torrentes no fueran adherentes, la corriente de agua sería prácticamente estéril, debido a que las bacterias se verían arrastradas aguas abajo con mucha mayor rapidez de la que ellas pudieran alcanzar nadando a contracorriente. Hemos establecido que un centímetro cuadrado de una superficie sumergida presenta frecuentemente un millón de bacterias, como mínimo, adheridas a ella, mientras que un centímetro cúbico del agua que discurre sobre dicha superficie contiene únicamente

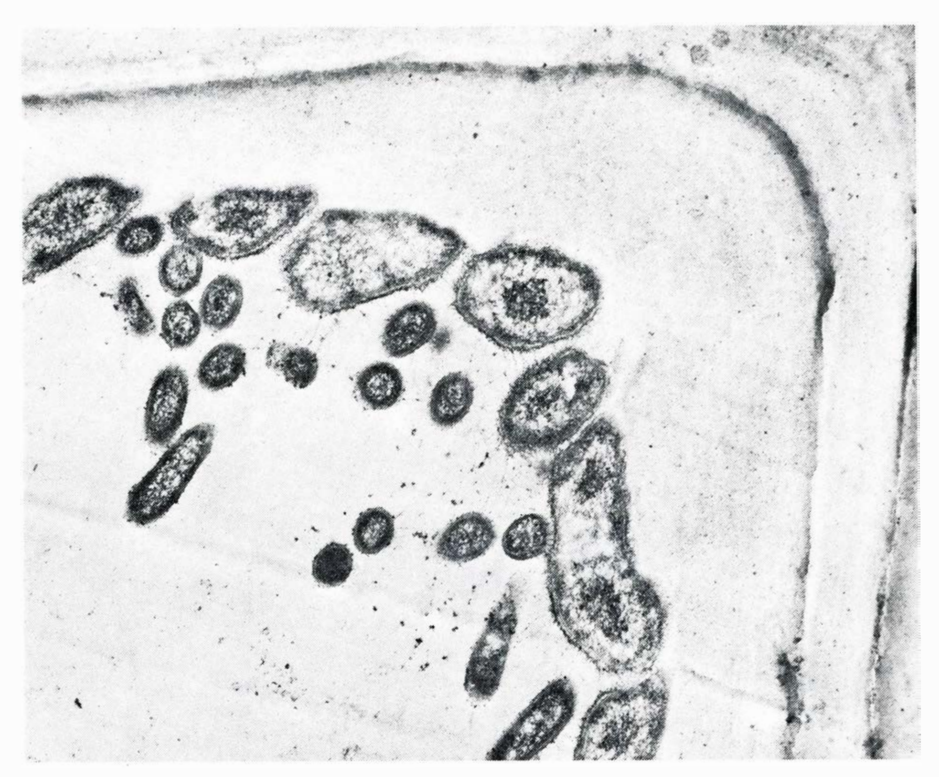


LA ADHERENCIA PATÒGENICA, por parte de una nueva clase de antibiótico, puede bloquearse, a fin de evitar o curar la infección. La adherencia de una bacteria *(arriba)* a una célula animal *(abajo)* por medio de enlaces covalentes o de lectinas *(a)* puede orillarse de tres formas distintas. Suministrando un análogo de los monómeros que constituyen las fibras del glicocálix bacteriano, que ocuparía los lados activos de los enzimas polimerizantes, evitándose así la síntesis de las fibras de polisacárido *(b)*. Los lados activos de una lectina pueden bloquearse por un análogo *(c)*. Los receptores de las glicoproteínas de las células animales pueden bloquearse por un agente que mimetice el glicocálix bacteriano *(d)*.

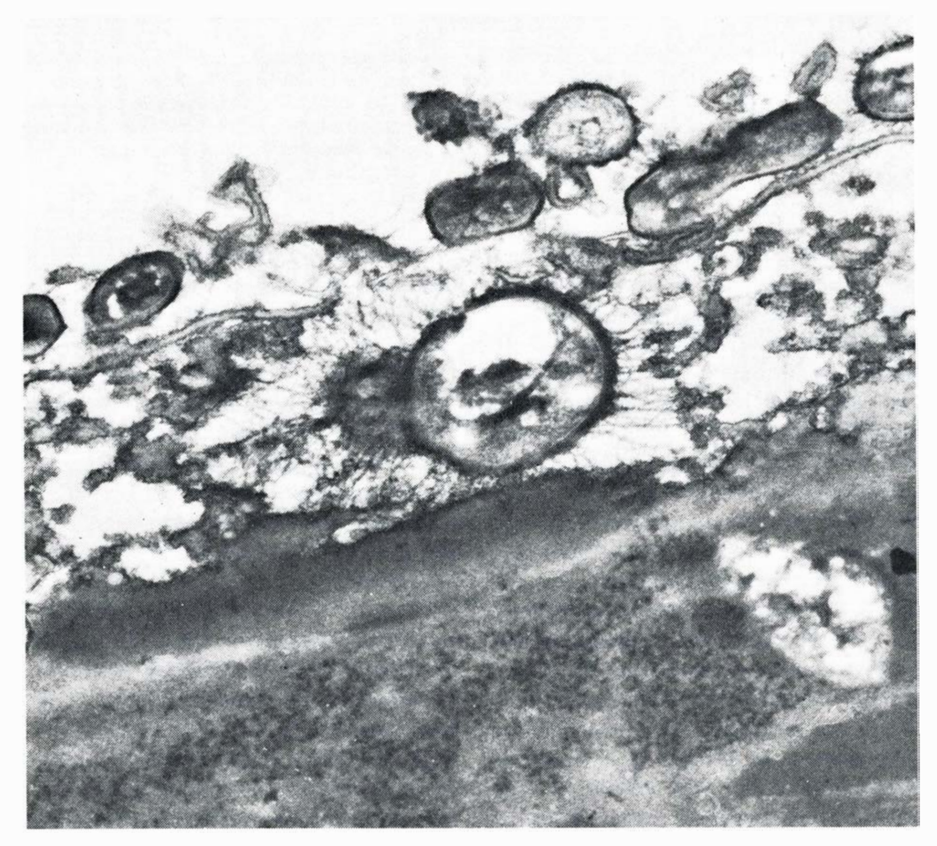
1000 bacterias. El valor adaptativo de la adherencia de las bacterias en estas circunstancias es comprensible. Las bacterias se nutren a expensas de los compuestos químicos que extraen del agua. La situación en una localización fija, con un aporte continuo de nutrientes, aireación vigorosa y excelente lavado de desechos proporcionados por la corriente de agua, favorece a las bacterias. En los torrentes polucionados, nosotros hemos hallado poblaciones que llegan a alcanzar los 10.000 millones de bacterias adheridas por cada centímetro cuadrado. (Estas poblaciones forman una gruesa película sobre las rocas del lecho del torrente, lo cual, por otra parte, favorece la eliminación de considerables cantidades de materia orgánica que poluciona el agua.)

La adherencia a un determinado tejido orgánico favorece a las bacterias que colonizan los tejidos animales o vegetales, del mismo modo que la adherencia a las rocas sumergidas en un torrente favorece a las bacterias que las habitan. Las células eucariotas suministran continuamente nutrientes orgánicos y presentan las condiciones físicas necesarias para el crecimiento bacteriano. Por ejemplo, Rolf Freter, de la Universidad de Michigan, ha descrito que las células de *Vibrio cholerae* se adhieren específicamente a los microvilli de las células epiteliales del intestino humano. Las células de *Neisseria gonorrhoeae* se adhieren a la mucosa de la uretra, y otros sistemas patógeno-huésped muestran el mismo comportamiento. La adherencia capacita a las bacterias infectantes para su permanencia en el lugar que colonizan, sobre todo si no está colonizado por otros microorganismos, como sucede en órganos normalmente estériles. Nosotros hemos comprobado que los patógenos persistentes del tracto urinario están provistos de un glicocálix bien desarrollado que impide el arrastre y eliminación por la orina de estas bacterias. La gran especificidad que algunos virus y bacterias tienen hacia determinados tejidos del huésped (por ejemplo, *Salmonella typhimurium* ataca a los folículos linfáticos del intestino y los poliovirus a las células nerviosas) puede ser explicada de forma adecuada por la especificidad del glicocálix de las células del tejido hospedante.

Entre otros ejemplos de especificidad citaremos a las bacterias digestoras del almidón de la panza de los bóvidos; éstas se adhieren tan específicamente a las partículas de almidón que la mejor manera para aislar dichas bacterias a partir de la población mixta de la panza consiste en recoger las partículas de almidón de los fluidos de la panza. A su vez, las bacterias bioluminiscentes que colonizan los órganos luminosos de los peces



BACTERIAS DIGESTORAS de la celulosa *(células de mayor tamaño).* **Se adhieren a la pared celular de una célula vegetal presente en el alimento. Una población de bacterias de menor tamaño se adhiere a las mayores, posiblemente a través de una interacción especifica entre los glicocálices de ambos tipos bacterianos. Esto constituye un ejemplo de "consorcio", una asociación cooperativa en la cual una población bacteriana suministra un nutriente esencial para la segunda población.**



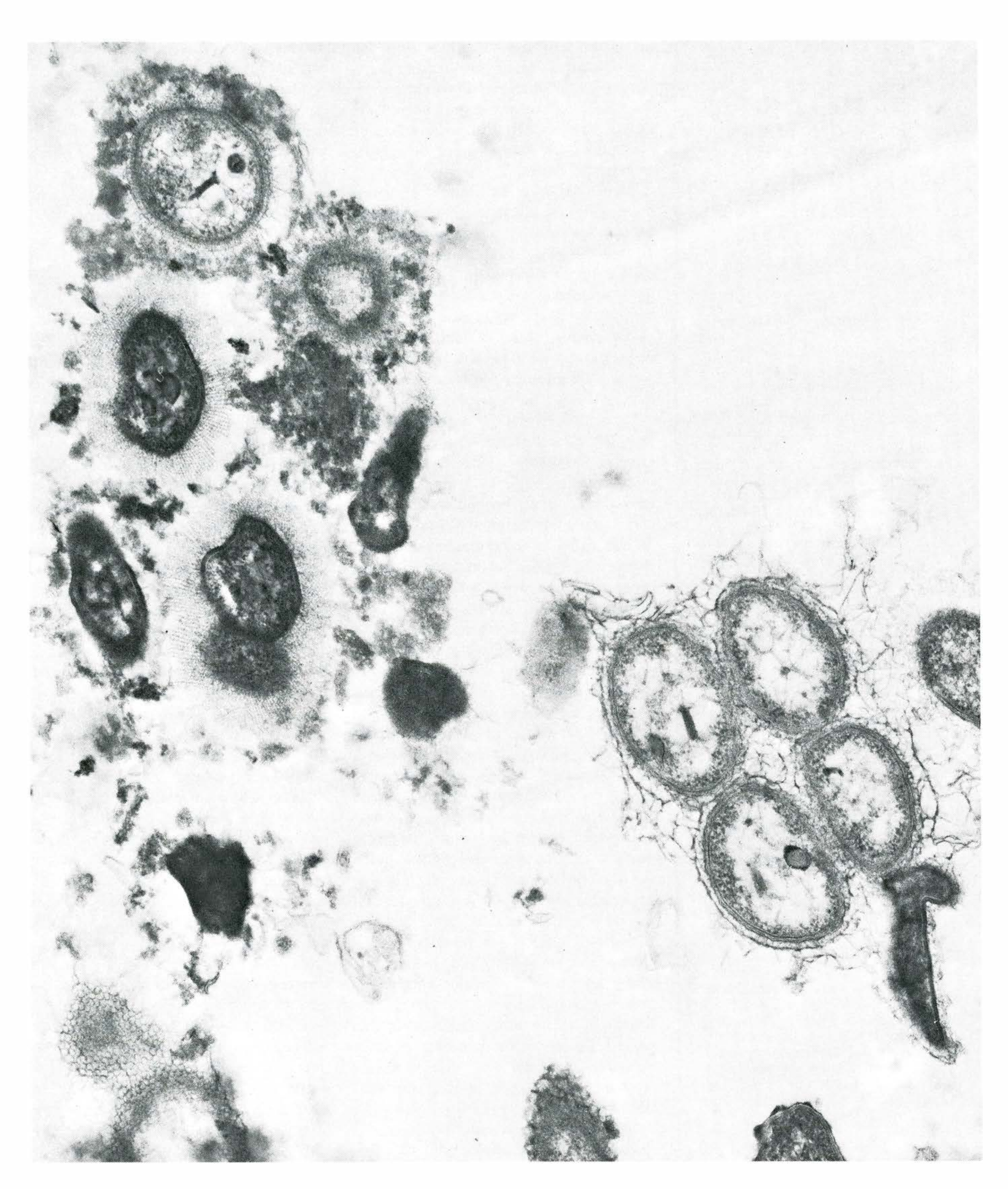
CELULA MUERTA, perteneciente a un tejido animal, que ha sido invadida por una bacteria adherente; ésta ha formado una cavidad a través de la membrana celular con su actividad digestora. Una población constituida por otras bacterias está adherida a la membrana de las célula animal.

de profundidad como *Photoblepharon,* se seleccionan entre la gran población de bacterias marinas por su capacidad de fijarse a las células que tapizan el órgano luminoso. En esta relación simbiótica, las bacterias proporcionan una fuente luminosa que utiliza el pez, y, a su vez, las bacterias se encuentran en un hábitat adecuado para su nutrición a expensas de las secreciones del pez.

Las fibras del glicocálix no sólo fijan a las bacterias, sino que preservan y concentran los enzimas líticos excretados por las mismas que, de este modo, ejercen una acción más directa contra la célula huésped. Las micrografías muestran cómo los enzimas producidos por una bacteria fijada a una partícula de alimento, como el material de las paredes celulares de los fragmentos de paja en la panza de un bóvido, "excavan" una cavidad dentro de la cual la bacteria realiza pausadamente su actividad metabólica. Las bacterias pueden colonizar las células de la piel y tejidos similares, antes de que sean renovadas por nuevas células, formando dichas cavidades. Esto sugiere que la problemática persistencia de *Staphylococcus aureus,* en la piel del personal sanitario, pueda ser debida a la dificultad de eliminar dichas bacterias de las cavidades digeridas por ellas en las células epiteliales de las que se nutren.

Un glicocálix puede funcionar también como un reservorio de alimentos para las bacterias. Como ya se ha dicho, las fibras de polisacárido están cargadas negativamente. Actuando de modo similar a las resinas de intercambio iónico, pueden fijar iones y moléculas de importancia metabólica que circulen por su entorno o que sean producidas por los enzimas líticos bacterianos, haciéndolos así más accesibles para las células.

En la naturaleza, el glicocálix ofrece una protección a las muchas situaciones traumáticas a que están sometidas las bacterias. La fijación a las superficies protege a las bacterias de determinados protozoos. Asimismo, constituye una barrera física contra las bacterias depredadoras y los bacteriófagos; y puede eliminar compuestos químicos nocivos para la célula, dada su capacidad fijadora de iones y moléculas. La capacidad de protección del glicocálix bacteriano tiene particular interés para las bacterias patógenas. Nosotros hemos observado que las células de *Pseudomonas aeruginosa,* desprovistas de glicocálix, introducidas en el pulmón de una rata, son rápidamente fagocitadas por los leucocitos de la rata; sin embargo, estos leucocitos son incapaces de fagocitar las microco-



FIBRAS DE GLICOCALIX, cuya gran variedad de modelos pueden observarse en la micrografía electrónica de estos cortes de diversas bacterias. Los glicocálices de esta población mixta de bacterias de la panza de un bóvido forman un mosaico de modelos de glicocálices designados como *(en el sentido de las agujas del reloj, a partir del ángulo inferior derecho):* desordenado, en panal de abeja, radial y radial-concéntrico. En la naturaleza, el glicocálix ofrece una protección a las muchas situaciones traumáticas a que están sometidas las bacterias. La fijación a las superficies protege a las bacterias de determinados protozoos. Asimismo, constituye una barrera física contra las bacterias depredadoras y los bacteriófagos; y puede eliminar compuestos químicos nocivos para la célula, dada su capacidad fijadora de iones y moléculas. La capacidad de protección del glicocálix bacteriano tiene particular interés para las bacterias patógenas. (Microfotografía que ha sido obtenida por J.W. Costerton.) El antibiótico mata a muchas de las bacterias no patógenas que se encuentran adheridas en condiciones normales y que ocupan muchos de los receptores disponibles, abriendo así una vía a la colonización de estos receptores por bacterias patógenas resistentes al antibiótico; estas cepas resistentes han sido seleccionadas por el prolongado tratamiento destinado a la "eliminación" del patógeno. La terapia "anti-glicocálix" hace innecesario mantener un largo tratamiento del enfermo con antibióticos convencionales. Urge, pues el desarrollo de nuevos métodos.

lonias de la misma bacteria si se hallan englobadas en un glicocálix. Otras experiencias ponen de manifiesto que el glicocálix impide que los anticuerpos del huésped reaccionen con la superficie de las células de *Pseudomonas aeruginosa* en el tracto urinario.

Parece ser que, en algunos procesos patológicos, la producción del glicocálix por las bacterias, y por consiguiente su virulencia, puede estimularse en presencia de concentraciones excesivas de ciertos nutrientes; éste puede ser el caso de la fibrosis cística, en la cual los pulmones afectados contienen concentraciones desmesuradas de iones sodio y magnesio. Algunos de nuestros resultados provisionales indican que los antibióticos no son capaces de evitar la capacidad de fijación del glicocálix bacteriano, y así lograr su objetivo. Como las bacterias infectantes, que causan muchas enfermedades, están probablemente agrupadas en microcolonias rodeadas por un glicocálix, es obvio que hay que tener en cuenta la penetración a través de las capas de polisacárido a la hora de diseñar los antibióticos.

Además de situar ventajosamente a las bacterias y protegerlas, el glicocálix puede agruparlas en una especie de comunidad organizada. En los sedimentos de los lagos y otros ecosistemas, dos o más especies bacterianas pueden actuar juntas, en un consorcio, para llevar a cabo procesos fisiológicos complejos. Ralph S. Wolfe, de la Universidad de Illinois en Urbana-Champaign, ha descrito consorcios en los cuales una especie bacteriana produce hidrógeno a partir de compuestos orgánicos; el hidrógeno es utilizado por otra especie bacteriana, para reducir el dióxido de carbono a metano (CH_4). Una transferencia fisiológica de esta naturaleza requiere que los miembros del consorcio estén en próxima vecindad, y esto sólo puede realizarse si existe una afinidad química entre los glicocálices de ambas especies celulares asociadas. En algunas de las micrografías parecen estar representados consorcios fisiológicos de esta naturaleza; estas micrografías muestran bacterias digiriendo la celulosa de una pared celular vegetal, mientras que bacterias de otra especie se colocan alrededor de las bacterias digestoras de celulosa.

Estas pruebas sueltas sugieren una estructura comunitaria en la que las células de una especie determinada se adhieren en un nicho ecológico favorable, próximo a la fuente del nutriente necesario. Estas poblaciones "adherentes" pueden responder a los cambios en las condiciones nutritivas o de otra índole con una plasticidad única. Las bacterias situadas en el nicho desfavorable mueren, dejando espacio y nutrientes a otras bacterias para las que el nicho alterado resulta ahora favorable. La población global mixta de células tendría así alguna de las ventajas de un organismo multicelular y ninguna de sus desventajas.

Dado que la adherencia parece desempeñar un importante papel en el éxito biológico de las bacterias patógenas, la prevención de la misma podría constituir un modo eficaz de combatir la infección bacteriana. Cabría la posibilidad de desarrollar una nueva clase de antibióticos que interfiriesen con la formación del glicocálix o su función en las bacterias patógenas. Existen como mínimo tres procedimientos mediante los cuales puede llevarse a cabo esta inhibición. El primero de ellos consistiría en bloquear la síntesis de las fibras de glicocálix. La polimerasa bacteriana que une los monosacáridos para formar las fibras sería inhibida si coexistiera con un compuesto que remedara el sustrato normal y, de este modo, ocupase el lado activo del enzima, no siendo, sin embargo, procesado para constituir la fibra normal de polisacárido. En ausencia de dichas fibras no habría glicocálix, ni adherencia, ni resistencia a los leucocitos. En segundo lugar, se podría encontrar un compuesto que pudiera ocupar y bloquear el lado activo de una lectina que mediatice la adherencia de las fibras del glicocálix bacteriano a las fibras de las células huésped.

Finalmente, sería posible bloquear los "receptores" en las células huésped, esto es, las fibras de glicoproteína a las que se adhieren las fibras bacterianas directamente. Un atractivo aspecto de un antibiótico dirigido contra el glicocálix es que no tendría que penetrar ni en la célula huésped ni en las células bacterianas, evitando de este modo dos problemas frecuentemente planteados por la terapia antibiótica: la toxicidad para las células huésped y la inducción de la resistencia bacteriana basada en los cambios de permeabilidad de la membrana celular bacteriana.

Quizás el mejor de los enfoques anteriormente mencionados sea el último de ellos. Como los polisacáridos son cadenas de azúcares más bien sencillas, y dado que la adherencia específica depende de la afinidad química existente entre los polisacáridos bacterianos y los de la célula huésped, la adherencia podría ser bloqueada mediante fragmentos del

material constituyente del glicocálix, por las subunidades de monosacárido o por análogos químicos de dichos compuestos. En este sentido, Onderdonk ha logrado establecer que tratando con material del glicocálix de *Bacteroides fragilis* el tejido peritoneal del ratón, se evita la fijación de esta bacteria al peritoneo; Onderdonk atribuye este efecto al bloqueo de los receptores de las células del ratón por el polisacárido del glicocálix bacteriano. Otros investigadores han descrito efectos similares en el buche del pollo. Este órgano está colonizado por una capa de lactobacilos. Tratando el buche de polluelos libres de gérmenes con polisacáridos del glicocálix de lactobacilos, se evita la formación de la película bacteriana.

Probablemente, las células de la garganta humana poseen receptores para los estreptococos, que causan las anginas y sus más graves efectos secundarios, incluyendo el reuma cardiaco. Estos receptores podrían bloquearse con monosacáridos análogos a los que constituyen el glicocálix de los estreptococos. La tarea de identificar estos polisacáridos y encontrar los análogos correspondientes podría ser particularmente útil en el caso de las infecciones estreptocócicas y en otras enfermedades en las que sucesivas infecciones pueden causar graves efectos. Prácticamente, en este tipo de infecciones se evitan las recidivas mediante un largo tratamiento con antibióticos convencionales, lo que puede ser peligroso. El antibiótico mata a muchas de las bacterias no patógenas que se encuentran adheridas en condiciones normales y que ocupan muchos de los receptores disponibles, abriendo así una vía a la colonización de estos receptores por bacterias patógenas resistentes al antibiótico; estas cepas resistentes han sido seleccionadas por el prolongado tratamiento destinado a la "eliminación" del patógeno. La terapia "antiglicocálix" hace innecesario mantener un largo tratamiento del enfermo con antibióticos convencionales.

El desarrollo de nuevos métodos que controlen las infecciones bacterianas requerirá, sin duda, un conocimiento mucho más detallado del que disponemos en la actualidad acerca de los constituyentes de los patógenos, en concreto y de sus células huésped. La realización de este objetivo puede llevar mucho tiempo. El concepto de la tenaz y específica adherencia bacteriana en relación con el ambiente que colonizan constituye un primer paso hacia un control y una manipulación más eficaz de las bacterias.

La eficiencia de los algoritmos

Los únicos métodos conocidos para resolver ciertos problemas matemáticos son demasiado lentos, incluso con los ordenadores más veloces. No se han descubierto otros métodos más eficientes, pero tampoco se ha demostrado que no existan

Harry R. Lewis y Christos H. Papadimitriou

Imagine el lector que se le pidiera preparar el itinerario de un representante de comercio que ha de visitar cierto número de ciudades. Para ello se le entrega un mapa en el que están señaladas las distancias entre las distintas ciudades, y se le pide que determine el circuito más corto que pase por todas las ciudades y regrese al punto de partida. Un método seguro de obtener la respuesta correcta consiste en trazar todos los posibles recorridos, medir sus longitudes y tomar el mínimo. Sin embargo, en cuanto el itinerario haya de contener cierto número de ciudades, habrá que comprobar, a buen seguro, centenares o millares de rutas posibles. Si fuesen 100 ciudades, por ejemplo, incluso el más veloz de los ordenadores necesitaría varias semanas de operación para hallar el recorrido mínimo.

Tratando de obtener más rápidamente la solución se podrían ensayar métodos menos rigurosos. Una idea de apariencia razonable sería visitar las ciudades más cercanas, antes de dirigirse a las alejadas. Pronto se descubre, sin embargo, que este procedimiento no siempre da el resultado apetecido. Otros posibles atajos resultan también fallidos. De hecho, los mejores métodos disponibles para resolver este problema no son mucho mejores que el evidente, aunque laborioso, procedimiento consistente en examinar todas las posibilidades. Los matemáticos sospechan en la actualidad que este problema, y otros muchos semejantes, pueden rebasar muy probablemente nuestra capacidad para resolverlos de alguna forma eficiente. No obstante, tales especulaciones tampoco han recibido confirmación, ni se ha podido demostrar que no existan métodos más rápidos.

En el problema del itinerario del representante de comercio, lo que tiene importancia no es la solución para un conjunto dado de ciudades concretas, sino el método general de hallar la solución para cualesquiera ciudades. Tales métodos se llaman algoritmos; se trata de procedimientos enunciados con la mayor precisión que dan un conjunto de reglas, aplicables siempre de igual forma a todos los casos concretos de un problema general. Cuando el problema ha de resolverse con auxilio de un ordenador, es indispensable disponer de un algoritmo, porque al ordenador solamente pueden serle presentados aquellos procedimientos que pueden enunciarse en la forma explícita e inambigua de algoritmos. Instrucciones vagas o que apelen a la intuición no son aceptables para el cómputo mecánico.

El método que se explica en la escuela para restar dos números enteros proporciona un ejemplo de algoritmo. Si se aplica correctamente cada uno de los pasos del método, el algoritmo dará siempre la solución correcta. Lo que es más importante, una vez aprendido el algoritmo, o almacenado en la memoria de un computador, o incorporado en la circuitería de una calculadora electrónica, puede aplicarse a un conjunto infinito de problemas de sustracción. Con este algoritmo puede determinarse siempre la diferencia entre dos números enteros cualesquiera.

En principio, todo problema para el que puede diseñarse un algoritmo puede resolverse mecánicamente. Por ello, quizá resulte sorprendente que haya problemas para los que existen algoritmos, pero que hasta el momento es imposible resolver en la práctica de forma general. Los algoritmos para resolver estos problemas darían siempre soluciones correctas, pero muy frecuentemente exigirían tiempos desorbitados. El problema del itinerario del viajante figura entre tales ímprobas tareas.

La eficiencia de los algoritmos de ordenador es tema de indudable importancia práctica. Tema que también tiene interés e importancia en otras ramas más formales de la matemática. Tanto en la matemática como en la lógica existen problemas para los que nunca podrán escribirse algoritmos. Por otra parte, se conocen ya algoritmos rápidos y eficientes para muchos problemas importantes. Comprendida entre ambos grupos se encuentra una tercera categoría de problemas que, en principio, siempre es posible resolver, para los cuales, sin embargo, sólo se conocen algoritmos ineficientes (y, por consiguiente, la mayoría de las veces, impracticables). Los matemáticos han podido demostrar que, para algunos de estos difíciles problemas, nunca podrán prepararse algoritmos eficientes. Para muchos de los problemas más importantes, sin embargo, se tiene únicamente la sospecha de que será imposible encontrar algoritmos eficientes.

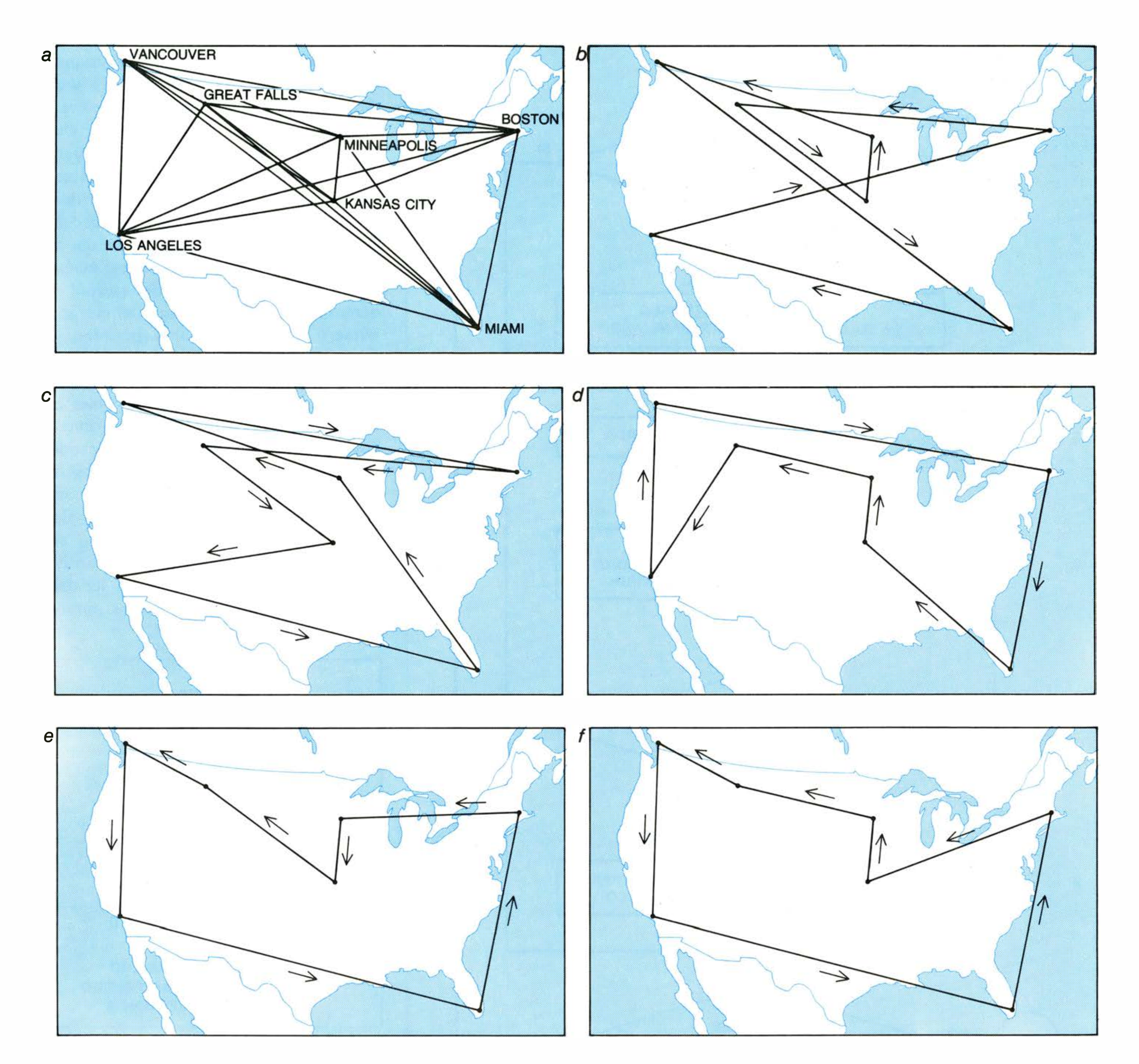
Puede haber más de un algoritmo que resuelva un problema dado. Por ejemplo, en Europa los niños aprenden a restar de forma ligeramente distinta que en los Estados Unidos. No obstante, ambos algoritmos dan resultados iguales y requieren el mismo tiempo de ejecución. Pero no ocurre invariablemente así con los distintos algoritmos que resuelven un mismo problema. Un famoso problema que puede resolverse bien por un algoritmo "rápido", bien por un algoritmo "lento", es el problema de los puentes de Königsberg.

En el siglo XVIII, en la entonces alemana ciudad de Königsberg (que hoy es la rusa Kaliningrado) se construyó un parque que ocupaba las orillas y dos islotes del río Pregel. Dentro del parque, siete puentes conectaban los islotes y las riberas. El problema de recorrer los siete puentes, pasando por cada uno de ellos una y solamente una vez, llegó a ser muy popular en la época.

Para resolver el problema carecen de importancia la forma y tamaño de los

islotes o la longitud de los puentes; la única información esencial es la configuración que definen las interconexiones. Se puede presentar matemáticamente tal información de forma muy concisa mediante un grafo, que es una estructura matemática constituida por un conjunto de puntos y de líneas que los enlazan. En el caso del parque de Königsberg, cada ribera y cada isla se condensan en un único punto, y cada puente se convierte en una línea que une dos puntos. Así pues, el grafo está formado por cuatro puntos y siete líneas. Si cada línea recibe un distintivo, cualquier recorrido por el parque queda completamente determinado por la sucesión de distintivos correspondiente.

Un procedimiento obvio de resolver el problema sería enumerar todos los caminos que cruzan todos los puentes



EL PROBLEMA DEL ITINERARIO de un representante de comercio es típico, y sirve para mostrar algunas graves deficiencias de los métodos disponibles en matemáticas. El problema puede resolverse, pero por métodos tan arduos y prolongados que, por lo general, no es práctico hallar la solución de esta forma. El problema es, dado un mapa en el que figuran los enlaces aéreos entre varias ciudades *(a)*, hallar un itinerario que determine la ruta mínima que pase por todas ellas y regrese al punto de partida. Un método que siempre daría resultado sería trazar todas las rutas posibles, medir sus longitudes y elegir la más corta. Sin embargo, aun con solamente siete ciudades hay más de 350 itinerarios posibles, y este número crece rápidamente (con la factorial del número de ciudades). Podrían ensayarse métodos abreviados. Hacer que el viajante se dirija desde el punto en que se encuentre al más alejado posible da, evidentemente, malos resultados; el itinerario así obtenido *(b)* se encuentra entre los más largos. Aún recorriendo las ciudades de una manera caprichosa, como, por ejemplo, en orden alfabético, se obtendrían mejores resultados *(c)*. Si se divide el país en segmentos, y visitando primero el Este, luego el Medio Oeste, y por último el Oeste, se obtiene un recorrido todavía más breve *(d)*. Finalmente, el representante podría viajar siempre a la ciudad no visitada más cercana a la que se encuentre. Por este procedimiento, se obtendrían dos itinerarios razonables, cuya diferencia depende de que se comience en Minneapolis *(e)* o en Kansas City *(f)*. Ni siquiera estas rutas son las mínimas posibles. (La solución se muestra en la ilustración de la página 89.) El tipo de solución que se busca es un algoritmo, un conjunto de instrucciones, que determine la ruta mínima para todo conjunto de ciudades.

y desechar después todos aquéllos que crucen un mismo puente más de una vez. Se trata de la técnica de búsqueda exhaustiva, análoga a la empleada en el problema del representante. Cuando le plantearon el problema de los puentes de Königsberg al matemático Leonhard Euler, éste reconoció inmediatamente la pobreza del método y descubrió otro. En reconocimiento a su trabajo, en la actualidad se llaman caminos eulerianos a los caminos que pasan por cada línea de un grafo una y exactamente una vez.

Euler escribió: "El problema particular de los siete puentes de Königsberg podría resolverse tabulando cuidadosamente todos los posibles recorridos, comprobando después por inspección cuáles, si es que existe alguno, satisfacen la condición. Este método de resolución, no obstante, es demasiado tedioso y demasiado difícil, en virtud del elevado número de combinaciones posibles, y, además, en otros problemas con muchos más puentes, sería imposible utilizarlo".

El otro método hallado por Euler es mucho más sencillo. Demostró primero que tendría que existir un recorrido como el deseado si el grafo cumpliera dos requisitos. Primero, que ha de ser posible ir desde cualquier punto del grafo hasta cualquier otro punto siguiendo las líneas del grafo; dicho de otra forma, el grafo no puede ser disconexo. Segundo, que todos los puntos

1 COMENZAR POR LA COLUMNA SITUADA A LA DERECHA

2 RESTAR A LA CIFRA SUPERIOR LA INFERIOR, USANDO LA TABLA. ANOTAR DEBAJO EL RESULTADO

3 ¿ES LA CIFRA SUPERIOR MAYOR O IGUAL QUE LA INFERIOR? SI / NO

4 SEÑALAR CON UN ASTERISCO LA COLUMNA QUE SE ESTA MANEJANDO

5 IR UNA COLUMNA HACIA LA IZQUIERDA

6 RESTARLE 1 A LA CIFRA SUPERIOR, USANDO LA TABLA. SUBSTITUIR EL DIGITO SUPERIOR POR EL VALOR OBTENIDO

7 ¿ES IGUAL A 9 EL VALOR DEL NUEVO DIGITO SUPERIOR? SI / NO

8 IR HACIA LA DERECHA HASTA ENCONTRAR COLUMNA SEÑALADA CON ASTERISCO

9 ¿QUEDAN DIGITOS SITUADOS A LA IZQUIERDA? NO → FIN / SI

10 IR UNA COLUMNA HACIA LA IZQUIERDA



MINUENDO

SUBSTRAENDO	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	9	0	1	2	3	4	5	6	7	8
2	8	9	0	1	2	3	4	5	6	7
3	7	8	9	0	1	2	3	4	5	6
4	6	7	8	9	0	1	2	3	4	5
5	5	6	7	8	9	0	1	2	3	4
6	4	5	6	7	8	9	0	1	2	3
7	3	4	5	6	7	8	9	0	1	2
8	2	3	4	5	6	7	8	9	0	1
9	1	2	3	4	5	6	7	8	9	0

2 * 3 *
3 0 4 7 MINUENDO
− 1 4 3 8 SUBSTRAENDO
1 6 0 9 DIFERENCIA

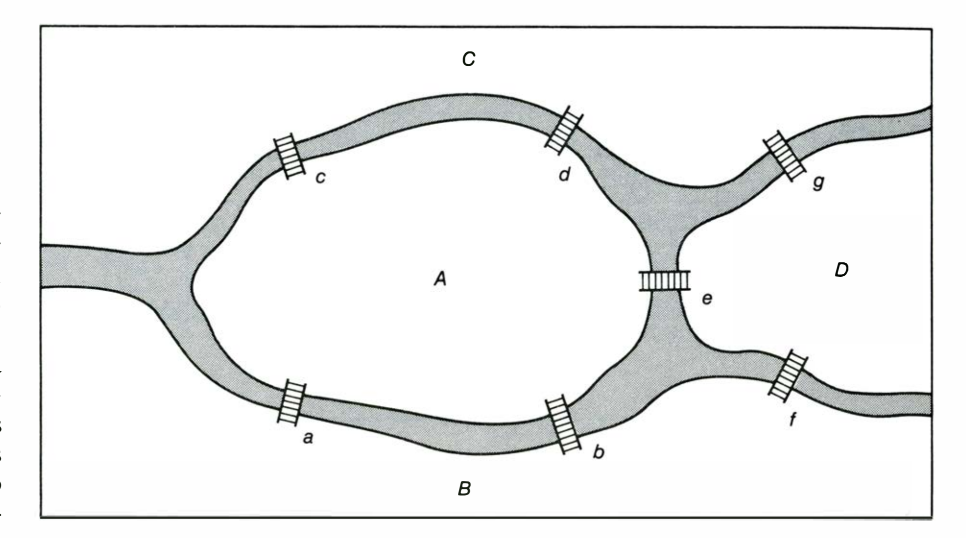
EL ALGORITMO de sustracción de números enteros determina un procedimiento explícitamente enunciado que puede aplicarse sin apelar a la intuición y sin necesidad de comprender la relación entre cada paso y la operación como un todo. El algoritmo, en el que se supone incorporada la tabla de diferencias que aquí se muestra, puede aplicarse a un número infinito de problemas de sustracción. Hay otros algoritmos igualmente efectivos para resolverlo. Por ejemplo, en Europa se enseña un método para restar ligeramente distinto del estadounidense: cuando se "lleva 1", éste se suma al dígito del substraendo en lugar de restárselo al dígito del minuendo.

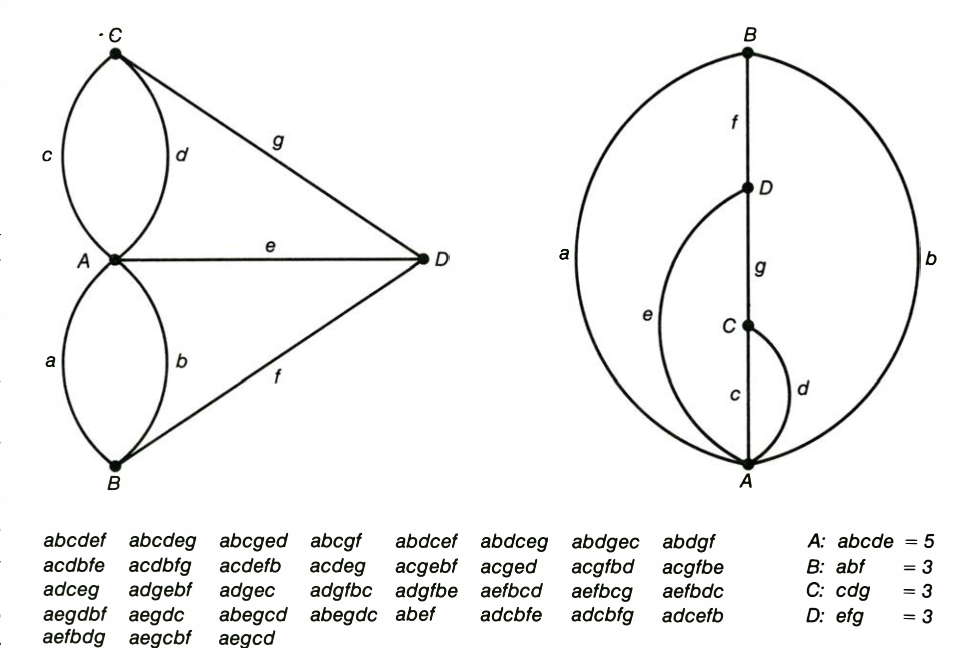
del grafo, a excepción de posiblemente dos, han de ser nudos donde se reúna un número par de líneas.

No es difícil comprender por qué un grafo no puede tener un recorrido euleriano si no cumple estas dos condiciones. Si ha de existir un camino que pase por todas las líneas del grafo, éste ha de ser conexo, es decir, todas sus regiones han de estar conectadas entre sí. Cada punto ha de tener un número par de líneas, pues la mitad de ellas son necesarias para llegar al punto, y, la otra mitad, para salir de él. Puede haber dos puntos con un número impar de líneas, si tales puntos se toman como origen y término del recorrido. La demostración de que todo grafo que reúna estos dos requisitos tiene efectivamente un camino euleriano requiere un razonamiento algo más complicado, que no ofreceremos aquí, pero que Euler construyó rigurosamente.

Es cosa sencilla expresar algorítmicamente la solución que Euler dio al problema anterior, que puede ser ejecutada mediante ordenador. Puede comprobarse que se satisface el primer requisito, conexión del grafo, marcando un punto cualquiera del mismo, y señalando después de la misma manera todos los puntos conectados mediante líneas con él. Se procede ahora de la misma forma a partir de los puntos recién marcados, y así sucesivamente. El grafo es conexo si al final se han marcado todos sus puntos. La segunda condición se verifica con igual facilidad. Se ordena a la máquina que examine cada punto del grafo y que cuente el número de líneas que concurren en ese punto. Si no hay más de dos puntos que tengan un número impar de líneas, el grafo tiene un recorrido euleriano. El parque de Königsberg cumplía la primera condición, pero no la segunda, y, por consiguiente, no existía ningún recorrido euleriano de sus siete puentes.

De los dos métodos, el de Euler es incuestionablemente el más económico para resolver el problema de los puentes de Königsberg; en él solamente es preciso enumerar cada punto y cada línea del grafo una vez, mientras que el método de búsqueda exhaustiva no termina hasta que se hayan enumerado todos los caminos que pasan por todos los puentes. El número de tales caminos es mucho mayor que el número de puntos y líneas del grafo. En este sentido el método de Euler da un algoritmo superior, pero ¿cuánto mejor? ¿Cómo puede medirse la diferencia entre ambos, y cómo puede saberse si la diferencia es significativa?



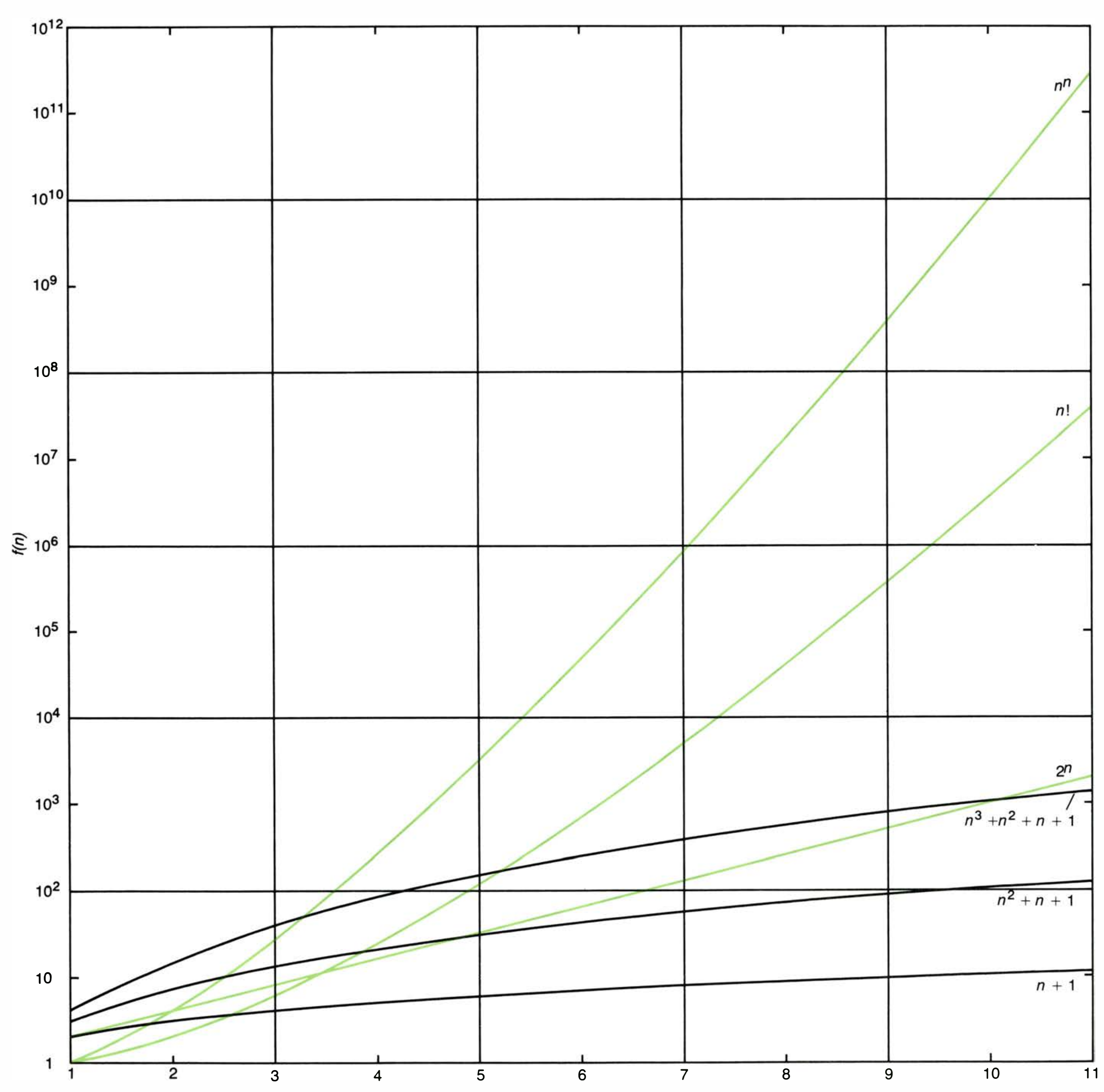


EL PROBLEMA DE EULER pide averiguar si en un grafo existe un cierto camino que pase exactamente una vez por cada línea. En este contexto, un grafo es, por definición, un conjunto de puntos juntamente con líneas que los conectan entre sí. El problema se planteó por primera vez respecto a la posibilidad de efectuar paseando un cierto recorrido de un parque existente en el siglo XVIII en la entonces ciudad alemana de Königsberg *(arriba)*; el camino pedido había de pasar exactamente una vez por cada uno de los siete puentes del parque. El parque puede representarse mediante un grafo al menos de dos formas equivalentes. Un método de resolver el problema sería hacer inventario completo de todos los caminos que van recorriendo el parque, avanzando todo lo posible sin repetir ninguna línea. Sin embargo, aún para un grafo pequeño hay muchos caminos así; en la lista que se da como ejemplo solamente están los caminos que comienzan por la línea *a*. Leonhard Euler descubrió un algoritmo más eficiente. Euler pudo demostrar que para que un grafo admita un recorrido que pase una sola vez por cada una de sus líneas es necesario y suficiente que en todos sus puntos (con la posible excepción de dos) concurran un número par de líneas. Tales recorridos se llaman caminos eulerianos. Contando las líneas que concurren en los puntos del grafo del parque de Königsberg, arriba esbozado, se comprueba que no existen recorridos eulerianos en él.

Para un grafo que tenga solamente cuatro puntos y siete líneas ambas técnicas son lo suficientemente rápidas como para ser consideradas practicables. Sin embargo, supongamos que se añaden al parque nuevas islas y puentes, o sea, que se añaden al grafo nuevos puntos y nuevas líneas. Si el problema ha de resolverse por el método de Euler, cada nuevo punto añade solamente un nombre a la lista de puntos a examinar. Por otra parte, si han de examinarse los posibles caminos mediante exploración exhaustiva, con cada nuevo punto o cada nueva línea, el tamaño de la lista queda multiplicado por un cierto factor. Un moderado aumento del tamaño del grafo provoca el crecimiento explosivo del número de caminos. La longitud de la lista de caminos acaba haciendo prohibitivo su examen.

La anterior comparación de las dos

FUNCION	TIPO	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$	$n = 11$
$f(n) = n + 1$	POLINOMICA	2	3	4	5	6	7	8	9	10	11	12
$f(n) = n^2 + n + 1$	POLINOMICA	3	7	13	21	31	43	57	73	91	111	133
$f(n) = n^3 \ n^2 + n + 1$	POLINOMICA	4	15	40	85	156	259	400	585	820	1,111	1,464
$f(n) = 2^n$	EXPONENCIAL	2	4	8	16	32	64	128	256	512	1,024	2,048
$f(n) = n!$	EXPONENCIAL	1	2	6	24	120	720	5,040	40,320	362,880	3,628,800	39,916,800
$f(n) = n^n$	EXPONENCIAL	1	4	27	256	3,125	46,656	823,543	16,777,216	387,420,489	10,000,000,000	285,311,670,611



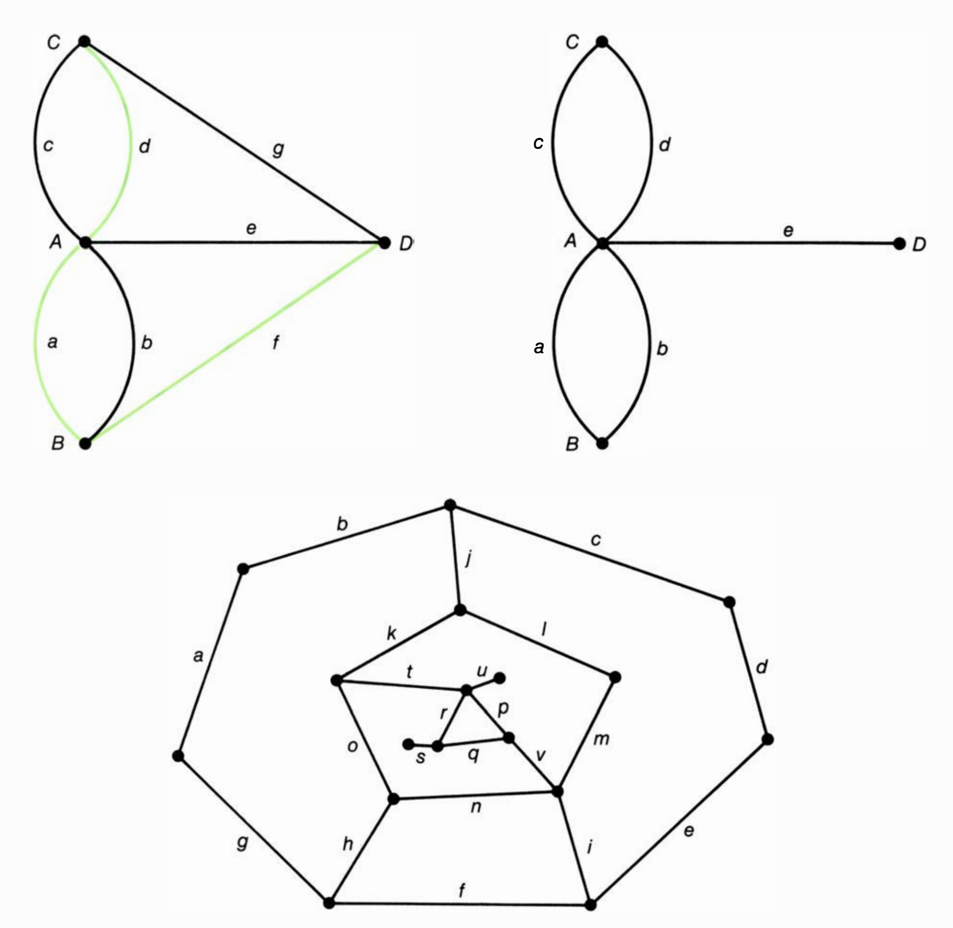
LA RAZON DE CRECIMIENTO del tiempo de ejecución de un algoritmo conforme aumenta el tamaño del problema se utiliza como criterio de viabilidad del algoritmo. Estas razones de crecimiento pueden describirse mediante funciones matemáticas, que pueden clasificarse como de tipo polinómico y de tipo exponencial. El valor de n representa cierta medida del tamaño del problema; los valores de las funciones corresponden a los tiempos de ejecución. Cualesquiera que sean la función exponencial y el polinomio que se tomen, siempre existe un valor de n a partir del cual la función exponencial siempre es mayor. Por esta razón, los algoritmos cuyo tiempo de ejecución crece como función exponencial del tamaño del problema se consideran ineficientes y carentes de valor práctico. En contraste con los anteriores, los algoritmos cuyo tiempo de ejecución crece en función polinómica son eficientes. En el problema de los puentes de Königsberg, el tiempo de ejecución con el método de Euler crece como la función polinómica $n + 1$; el método de tabulación exhaustiva exige un tiempo dado por una función exponencial de la forma 2^n.

soluciones del problema de Euler proporciona las bases para diseñar un método completamente general de evaluar la rapidez o la viabilidad de un algoritmo, esto es, de su eficiencia. Se imagina que al algoritmo se le suministran entradas cada vez mayores, y se observa la razón de crecimiento del tiempo invertido en ejecutarlo. De este modo, pueden hacerse juicios inequívocos sobre la calidad del algoritmo. El método de exploración exhaustiva no es solamente un método más lento; en general, es demasiado lento para tener valor práctico alguno. El método de Euler sigue siendo válido para problemas de tamaño, en esencia, ilimitado.

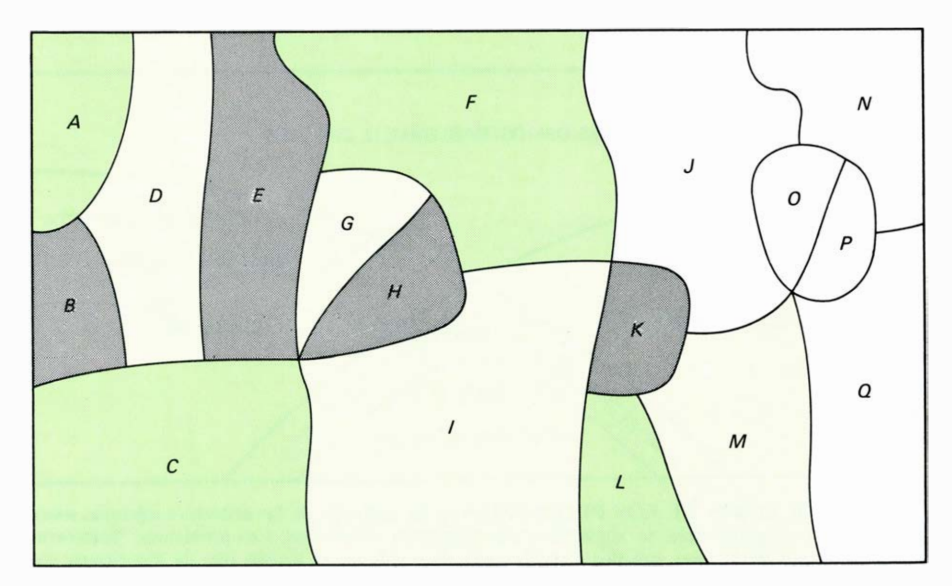
Conforme aumenta el tamaño de los grafos a examinar, el tamaño de las listas obtenidas por el método de exploración exhaustiva crece exponencialmente. Cada vez que se añada al grafo un número fijo de puntos y líneas se duplicará el tamaño de la lista resultante. Un crecimiento de esta naturaleza puede expresarse mediante una función matemática de la forma 2^n, siendo n cierta medida del tamaño y complejidad del grafo. Otras muchas funciones tienen tasas de crecimiento parecidas, o aún mayores. Entre ellas se tienen n^n y $n!$ (que se lee "factorial de n" y representa el producto de n por todos los enteros comprendidos entre 1 y n). Para los fines de esta discusión será suficiente considerar que todas estas funciones tienen la propiedad de crecimiento exponencial.

Otro tipo de funciones matemáticas son las llamadas funciones polinómicas. Las más sencillas de esta clase son las funciones lineales, como, por ejemplo, $3n$, que representan la simple relación de proporcionalidad. El tiempo necesario para resolver el problema de los puentes de Königsberg por el método de Euler crece en función lineal del tamaño del grafo. Otros polinomios son, por ejemplo, n^2, n^3, y sucesivos, y la suma de productos por números de tales funciones. Lo que distingue a los polinomios de las funciones exponenciales es que la n nunca aparece como exponente.

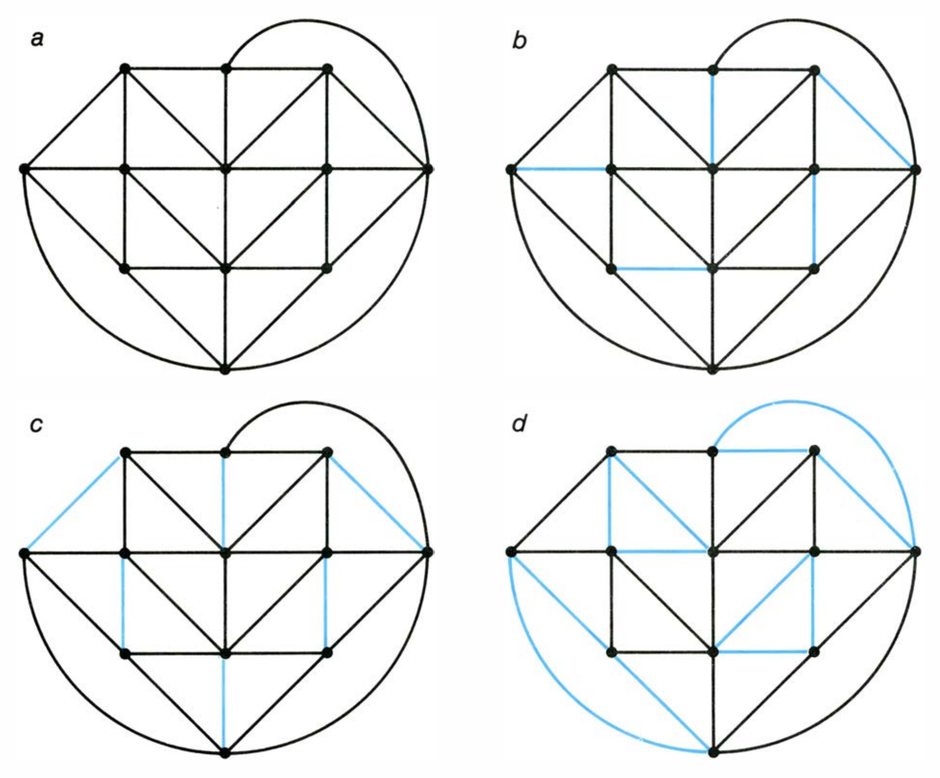
Tomando valores de n suficientemente grandes, toda función exponencial alcanzará y superará a cualquier función polinómica. Esta fue la causa del derrotismo de Thomas Malthus, al constatar la tasa exponencial de crecimiento de las poblaciones y compararlo con el crecimiento polinómico de la producción de alimentos. Para valores pequeños de n es perfectamente posible que una función polinómica dada sea mayor que una



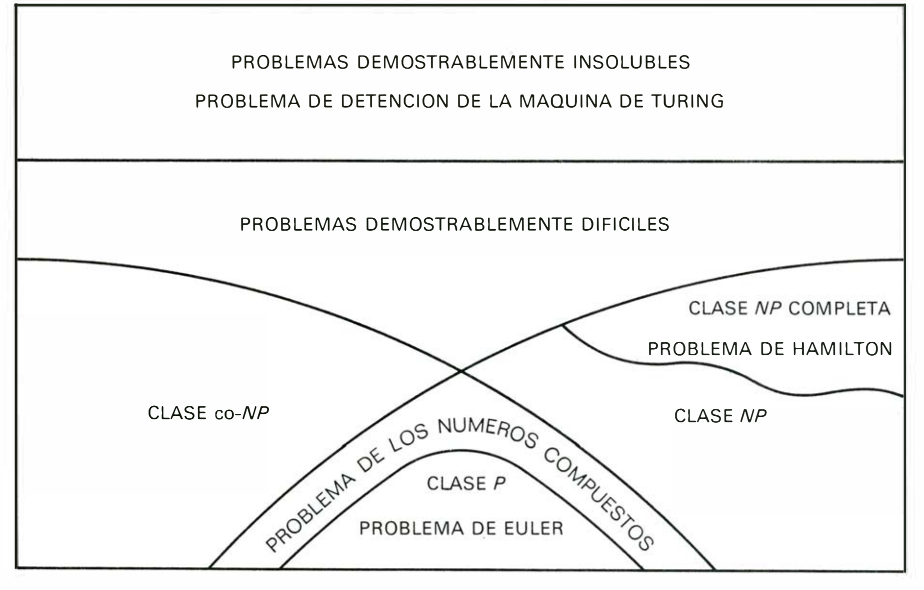
EL PROBLEMA DE HAMILTON, enunciado por el matemático irlandés William Rowan Hamilton, es, en apariencia, semejante al de Euler, aunque ahora se pide si existe un camino que pase una y solamente una vez por todos los puntos de un grafo (en lugar de las líneas). El grafo deducido del parque de Königsberg *(arriba, izquierda)* tiene un recorrido hamiltoniano *(en color)* a pesar de no tenerlo euleriano. Suprimiendo dos líneas se forma un grafo *(arriba, derecha)* que tiene un camino euleriano (que comienza en el punto *D* y prosigue, por ejemplo, a lo largo de las líneas *e, a, b, c* y *d)*, pero no tiene recorrido hamiltoniano. El tercer grafo *(abajo)* no tiene ninguna de las dos propiedades. No se conocen algoritmos que resuelvan eficientemente el problema de Hamilton; los métodos disponibles no son mucho mejores que la simple exploración exhaustiva.



EL PROBLEMA DEL MAPA TRICOLOR pregunta si es posible colorear con tres tonos las regiones de un mapa dado, de forma que nunca haya dos regiones fronterizas del mismo color. Para un mapa concreto, el problema puede resolverse, en teoría, examinando todas las posibles coloraciones, pero tal procedimiento es extraordinariamente tedioso; el mapa que se muestra tiene 3^{16}, que son alrededor de 43 millones, de coloreados posibles. No se conoce ningún algoritmo eficiente para resolver el problema, pero cuando existe una solución tricolor, tal solución podría hallarse, al menos en un principio, por tanteo (que no es un método algorítmico). Cuando no hay solución, los tanteos no tienen ningún valor. La solución ensayada falla al llegar a la región *J*, que comunica con los otros tres colores y, por consiguiente, no puede recibir ninguno de ellos. Tal fracaso no prueba que no exista ninguna solución tricolor; se puede encontrar una solución intercambiando los colores de las regiones *K* y *L*, y continuando la serie de tanteos en el resto del grafo.



LA ASIGNACION DE HABITACIONES en una residencia es problema que puede resolverse eficientemente si cada habitación ha de ser compartida por dos estudiantes, pero no, en cambio, si han de compartirla tres. El problema puede enunciarse mediante un grafo, en el que cada estudiante está representado por un punto y donde cada línea entre dos puntos significa que los correspondientes estudiantes serían compañeros de habitación compatibles *(a)*. La asignación de habitaciones a pares de estudiantes se efectúa mediante una técnica llamada de aumentación: se hace primero una asignación incompleta *(b)* y después se halla una completa (si existe) mediante una sucesión de arreglos *(c)*. No existe un método análogo eficiente para hallar grupos compatibles de tres estudiantes. Una vez hallado un reparto satisfactorio, puede exhibirse sin ninguna dificultad *(d)*.



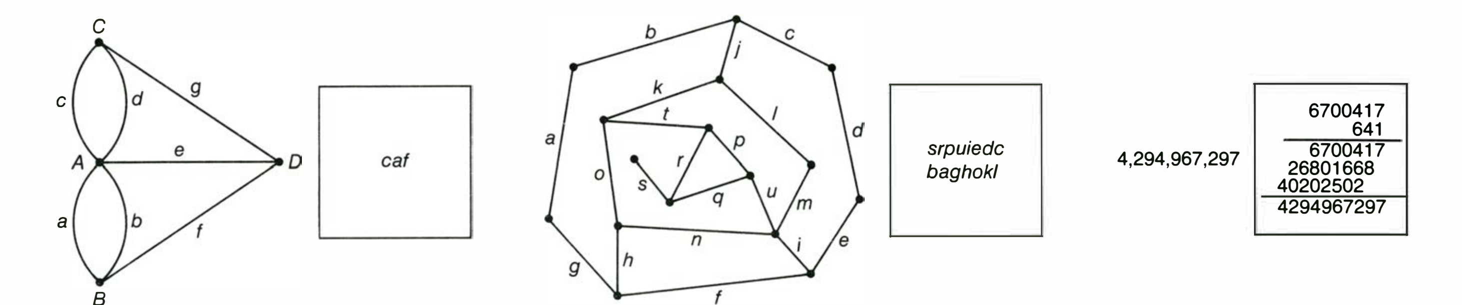
LA CLASIFICACION DE LOS PROBLEMAS se ha deducido de las actuales conjeturas sobre la existencia o inexistencia de algoritmos eficientes para resolverlos. Los problemas "demostrablemente irresolubles" son aquellos para los que no puede haber ningún tipo de algoritmos; los "demostrablemente difíciles" solamente tienen algoritmos de tiempo exponencial. Los problemas para los que se conocen algoritmos de tiempo polinómico se asignan a la clase *P* (inicial de polinomio). El carácter de los restantes problemas es menos seguro: solamente se conocen algoritmos suyos de tiempo exponencial, pero no está demostrada la inexistencia de algoritmos eficientes. Estos problemas se asignan a las clases *NP* y co-*NP*. Las letras *NP* significan "no determinísticamente polinómicos" y, con suerte, podrían resolverse rápidamente por tanteo. La clase co-*NP* contiene todos los problemas cuya solución es la disyuntiva "sí" o "no", y cuyas versiones complementarias "no" o "sí" pertenecen a *NP*. La clase *NP*-completa es un subconjunto de la clase *NP* compuesta con los problemas que tienen la siguiente propiedad especial: Si uno de ellos pudiera resolverse mediante un algoritmo eficiente, entonces todos los demás problemas de la clase *NP* podrían resolverse eficientemente. El descubrimiento de un tal algoritmo sería demostración de que *P* y *NP* son idénticas, pero la mayoría de los matemáticos creen que tal algoritmo no puede existir.

función exponencial dada, pero siempre existe un valor *n*, dependiente de las funciones particulares consideradas, a partir del cual la mayor de las dos es la función exponencial. La forma precisa que pueda tener el polinomio no importa gran cosa; únicamente cambiará el punto en el que la función polinómica se verá superada.

Al parecer, todos los especialistas en ciencias de la computación están de acuerdo actualmente en que los algoritmos cuyo tiempo de ejecución crece exponencialmente en función del tamaño de la entrada carecen de valor práctico. Llamaremos algoritmos de "tiempo exponencial" a los algoritmos de esta categoría, o, más sencillamente, algoritmos no eficientes. Los únicos algoritmos que se consideran lo suficientemente rápidos o eficientes para ser aplicables en general son los de "tiempo polinómico".

Como es obvio, incluso entre los algoritmos eficientes, unos son más rápidos que otros; ahora bien, para los objetivos de nuestra discusión tan sólo importa distinguir entre algoritmos de tiempo polinómico y algoritmos de tiempo exponencial, pues ambos forman categorías distintas. Este sistema de clasificación tiene además la ventaja de que la rapidez de un algoritmo es propiedad intrínseca del mismo, y es independiente de la máquina que lo ejecute. Haciendo suficientemente grande el tamaño del problema, un algoritmo de tiempo polinómico, aunque esté ejecutado en la más lenta de las máquinas, dará la respuesta antes que lo haga un algoritmo exponencial en el más veloz de los ordenadores.

Las propiedades matemáticas de los algoritmos fueron estudiadas en la década de 1930 por el matemático británico A. M. Turing, inventor de un ordenador imaginario hoy denominado "máquina de Turing". En su concepción, la máquina de Turing era un autómata provisto de una cinta infinita de papel dividido en regiones cuadradas. La máquina solamente podía realizar cuatro acciones: que la cinta avanzase un cuadro, poner una marca en un cuadro, borrar una marca ya existente en él y, al finalizar un cálculo, podía detenerse. Estas operaciones habían de realizarse de acuerdo con una sucesión de instrucciones que formaban parte del mecanismo interno. Evidentemente, Turing nunca llegó a construir tal máquina: se trataba sencillamente de un artificio conceptual para resolver de manera automática problemas de lógica y de matemática. Lo que, en realidad, interesaba a Turing



LOS PROBLEMAS DE CLASE *NP* plantean una disyuntiva, a contestar "sí" o "no", cuya elucidación suele requerir procedimientos lentos e ineficientes; pero cuando la respuesta es afirmativa, tal hecho puede comprobarse de forma breve. No existe por el momento ningún método eficiente para determinar si un grafo tiene un camino hamiltoniano, pero si lo tiene, se puede expedir un breve "certificado" que así lo demuestre. El certificado consiste, simplemente, en enumerar las líneas del camino en el mismo orden en que son recorridas. Otro problema de clase *NP* pide determinar si un número es compuesto, es decir, si puede expresarse como producto de otros dos. No se conoce ningún método eficiente para resolver esta cuestión; de hecho, en un caso particular se necesitaron más de 100 años para demostrar que un cierto número era compuesto. En 1640, Pierre de Fermat conjeturó que el número 4.294.967.297, que es igual a $2^{32} + 1$, es número primo. Su error perduró hasta 1732. Sin embargo, sabiendo que un número es compuesto, basta para comprobarlo una multiplicación cuyo producto resultante sea el número en cuestión.

no era la resolución efectiva de problemas mediante su máquina, sino determinar qué tipos de problemas podrían resolverse con ella, y cuáles no.

Turing descubrió que incluso una máquina tan sencilla como ésta podría resolver cualquier problema para el que se tuviese un algoritmo. El cálculo sería laborioso e indirecto, pero concediéndole suficiente tiempo y papel, la máquina terminaría por hallar la solución y detenerse. Reducida a sus elementos más esenciales, la máquina de Turing es un lenguaje para enunciar algoritmos, en principio tan potente como los más complejos lenguajes que se utilizan actualmente para comunicarnos con los ordenadores.

Además de concebir tales máquinas, Turing puso de manifiesto sus limitaciones. En 1936, demostró que hay problemas que no pueden resolverse mediante máquinas de Turing, y de aquí se sigue que tales problemas no podrán ser resueltos por ningún ordenador automático. Existen problemas para los que es imposible enunciar un algoritmo, ni siquiera en teoría. El primer ejemplo estudiado por Turing es el problema de predecir si una máquina de Turing concreta, una vez puesta en marcha, llegará a terminar su cálculo y se detendrá. Analizando este problema, llegó a la demostración de que no puede existir un procedimiento general para decidir si las proposiciones matemáticas son verdaderas o falsas. Desde entonces, se han propuesto diversidad de otros problemas con las mismas propiedades.

Un resultado del trabajo de Turing fue la división de todos los problemas matemáticos imaginables en dos clases. Los problemas para los que nunca podrán escribirse algoritmos son, desde el punto de vista formal, permanentemente irresolubles. Ciertos casos de tales problemas podrán resolverse si se tiene suerte o alguna clase de intuición genial, pero nunca podrá hallarse un método general de resolución de los mismos.

Todos los demás problemas de la matemática y la lógica pueden ser resueltos mediante algoritmos. Como ya hemos visto, unos algoritmos son más útiles que otros. La clase de los problemas resolubles puede, por tanto, descomponerse en dos subclases. Una, la formada por aquellos problemas que poseen algoritmos eficientes, de tiempo polinómico; la otra, compuesta por los problemas que solamente admiten algoritmos de tiempo exponencial. El problema de Euler es miembro de la clase de problemas con tiempo polinómico, pues el propio método de Euler es un algoritmo de tiempo polinómico. Se conocen también problemas que solamente pueden poseer algoritmos de tiempo exponencial, aunque se trata de problemas bastante obscuros.

A pesar de que estos dos grupos de problemas son bien distintos, no siempre es tarea fácil asignar a un problema dado el grupo al que pertenece. En realidad, hay una clase muy interesante de problemas que parecen ocupar una posición intermedia. Para estos problemas no se conocen algoritmos eficientes, y las mejores soluciones disponibles requieren tiempos exponencialmente crecientes con el tamaño del problema; pero nadie ha podido todavía demostrar que no tienen soluciones de tiempo polinómico.

Uno de tales problemas es el estudiado por William Rowan Hamilton, matemático irlandés del siglo pasado. En apariencia, el problema de Hamilton es muy semejante al de Euler. El problema de Hamilton consiste en determinar si un grafo dado tiene un recorrido que pase exactamente una vez por cada uno de sus puntos (mientras que en el de Euler se buscaba un camino que pasara por cada línea una sola vez). En realidad, ambos problemas son muy diferentes, y el método de Euler no es aplicable al problema de Hamilton. El grafo deducido del plano del parque de Königsberg resulta poseer un recorrido hamiltoniano, aunque, como ya hemos visto, no tiene ningún recorrido euleriano. Por otra parte, suprimiendo dos líneas se obtiene un grafo que tiene recorrido euleriano, pero ninguno hamiltoniano. Otros numerosos grafos carecen de recorridos de alguno de estos dos tipos.

El problema de Hamilton puede resolverse por exploración exhaustiva; de hecho, el procedimiento no difiere sustancialmente del empleado para enumerar todos los caminos posibles que pudieran poseer la propiedad euleriana. Sin embargo, para el problema de Hamilton no se ha encontrado ningún algoritmo eficiente comparable al del método de Euler. El problema recibió considerable atención por parte de muchos de los mejores matemáticos del siglo pasado, pero aún así, los mejores métodos de análisis del problema disponibles en la actualidad no son mucho mejores que la tabulación exhaustiva. Por otra parte, todos los intentos encaminados a demostrar que no existen métodos más eficientes han fracasado también, y ha de considerarse posible que se descubra un algoritmo eficiente en el día de mañana.

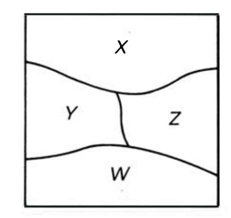
Los problemas para los que se sabe que poseen soluciones de tiempo polinómico, como, por ejemplo, el problema de Euler, se dice que pertenecen a la clase *P*

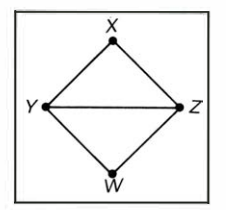
(inicial de "polinomio"). El problema de Hamilton pertenece a otra clase, que se designa por las siglas *NP* (que significan "nondeterministic polynomial", no determinísticamente polinómicos). La clase *NP* incluye todos los problemas de la clase *P;* dicho de otra forma, *P* es subconjunto de *NP.* Además de los anteriores, la clase *NP* contiene otros problemas de *status* menos seguro. Todos ellos son problemas resolubles al menos en teoría; tienen algoritmos, aunque por el momento solamente se conocen algoritmos de tiempo exponencial. Podrían tal vez tener algoritmos de tiempo polinómico (en cuyo caso *NP* y *P* serían idénticas) o pueden resultar permanentemente inabordables, y tener solamente soluciones ineficientes.

Los problemas aquí considerados, y todos los problemas clasificados de esta forma, pueden describirse como un conjunto infinito de cuestiones semejantes, cada una de las cuales puede responderse con "sí" o con "no". En el caso de problemas formalmente insolubles, como el problema de predecir si una máquina de Turing determinada se detendrá o no, estas preguntas no pueden contestarse mediante ningún procedimiento algorítmico. Para los problemas de la clase *P,* las cuestiones tienen respuesta siempre, tanto si la respuesta es afirmativa como si es negativa, por un procedimiento eficiente. Para que un problema forme parte de la clase *NP* no es preciso que exista un método eficiente de responder las preguntas afirmativa o negativamente. Lo que se requiere es que cuando la respuesta sea afirmativa exista un método breve y convincente que así lo demuestre.

El problema de Hamilton, por ejemplo, satisface este requisito. Ninguno de los métodos hoy conocidos permite decidir eficientemente si un grafo dado tiene recorrido de Hamilton, pero cuando lo tiene es posible mostrar el recorrido en cuestión, y demostrar fácilmente que el recorrido existe. Así pues, para todo grafo de Hamilton puede "extenderse un certificado" que demuestre que pertenece a esta clase especial de grafos. Un tal certificado consistiría en nombrar las líneas del grafo en el mismo orden en que las iría recorriendo el camino hamiltoniano. Para hallar tal camino, quizá sean precisas varias semanas de tabulaciones, pero una vez hallado puede mostrarse fácilmente. Otro problema perteneciente a la clase *NP* es la cuestión de ver si un número entero dado es primo o compuesto, esto es, si es o no descomponible en producto de otros dos números. Como antes ocurría, no se conoce ningún método eficiente de responder a esta cuestión; pero cuando el número es verdaderamente compuesto, tal hecho puede probarse muy brevemente, a saber, mediante una multiplicación correcta cuyo resultado sea el número propuesto.

Es preciso poner gran atención al formular las preguntas a responder con un sí o con un no en los problemas de la categoría *NP,* pues el problema no-o-sí complementario del dado puede no pertenecer a la misma clase. Por ejemplo, el complementario del problema de Hamilton pediría demostrar que un grafo no tiene ningún camino que pase una sola vez por cada uno de sus puntos. Este problema puede muy bien no pertenecer a la clase *NP,* pues, hasta el momento, la única forma de demostrar la no existencia de tal camino sería dar el inventario completo de todos los caminos posibles,

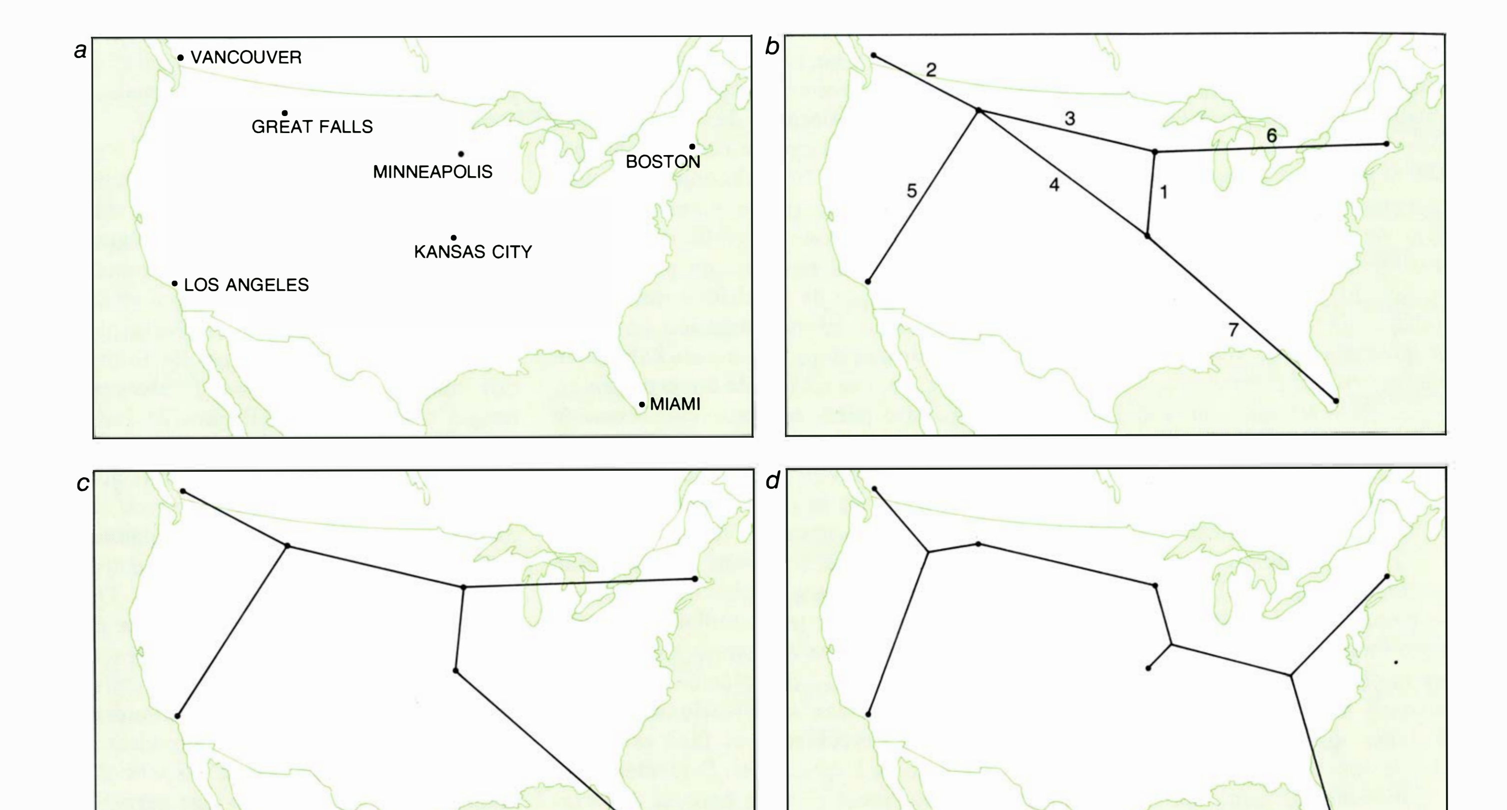




∧ Y
∨ O
¬ NO

[(*X* es rojo) ∨ (*X* es azul) ∨ (*X* es verde)] ∧ [(*Y* es rojo) ∨ (*Y* es azul) ∨ (*Y* es verde)] ∧ [(*W* es rojo) ∨ (*W* es azul) ∨ (*W* es verde)] ∧ [(*Z* es rojo) ∨ (*Z* es azul) ∨ (*Z* es verde)] ∧	Todo país recibe algún color
[(¬*X* es rojo) ∨ (¬*X* es azul)] ∧ [(¬*X* es rojo) ∨ (¬*X* es verde)] ∧ [(¬*X* es azul) ∨ (¬*X* es verde)] ∧ [(¬*Y* es rojo) ∨ (¬*Y* es azul)] ∧ [(¬*Y* es rojo) ∨ (¬*Y* es verde)] ∧ [(¬*Y* es azul) ∨ (¬*Y* es verde)] ∧ [(¬*W* es rojo) ∨ (¬*W* es azul)] ∧ [(¬*W* es rojo) ∨ (¬*W* es verde)] ∧ [(¬*W* es azul) ∨ (¬*W* es verde)] ∧ [(¬*Z* es rojo) ∨ (¬*Z* es azul)] ∧ [(¬*Z* es rojo) ∨ (¬*Z* es verde)] ∧ [(¬*Z* es azul) ∨ (¬*Z* es verde)] ∧	Ningún país recibe dos colores
[(¬*X* es rojo) ∨ (¬*Y* es rojo)] ∧ [(¬*X* es azul) ∨ (¬*Y* es azul)] ∧ [(¬*X* es verde) ∨ (¬*Y* es verde)] ∧	*X* e *Y* no tienen igual color
[(¬*X* es rojo) ∨ (¬*Z* es rojo)] ∧ [(¬*X* es azul) ∨ (¬*Z* es azul)] ∧ [(¬*X* es verde) ∨ (¬*Z* es verde)] ∧	*X* y *Z* no tienen igual color
[(¬*Y* es rojo) ∨ (¬*Z* es rojo)] ∧ [(¬*Y* es azul) ∨ (¬*Z* es azul)] ∧ [(¬*Y* es verde) ∨ (¬*Z* es verde)] ∧	*Y* y *Z* no tienen igual color
[(¬*Y* es rojo) ∨ (¬*W* es rojo)] ∧ [(¬*Y* es azul) ∨ (¬*W* es azul)] ∧ [(¬*Y* es verde) ∨ (¬*W* es verde)] ∧	*Y* y *W* no tienen igual color
[(¬*W* es rojo) ∨ (¬*Z* es rojo)] ∧ [(¬*W* es azul) ∨ (¬*Z* es azul)] ∧ [(¬*W* es verde) ∨ (¬*Z* es verde)]	*W* y *Z* no tienen igual color

EL CALCULO PROPOSICIONAL sirve de lenguaje universal para los problemas de clase *NP*. El problema que aquí hemos tomado consiste en colorear un mapa, o su grafo equivalente, con tres colores. Una frase, o fórmula, del cálculo proposicional está compuesta por enunciados, tales como "*X* es rojo", enlazados por las partículas conectivas lógicas "y", "o" y "no". Al unir dos enunciados mediante la conectiva "y", la fórmula resultante es verdadera solamente cuando ambos enunciados son verdaderos. Para ser verdadera una fórmula con la conectiva "o" es preciso que lo sea al menos uno de sus enunciados componentes. "No" requiere que el enunciado sea falso. La fórmula que representa el problema de colorear el mapa tiene tres componentes: los dos primeros establecen que cada país tiene exactamente un color, y la tercera parte especifica qué países no pueden tener colores iguales. El problema de colorear el mapa se traduce así en determinar si esta fórmula es satisfactible, es decir, si existe alguna forma de suponer verdaderos algunos de sus enunciados y falsos los restantes, sin que aparezcan contradicciones. Como todo problema *NP* puede expresarse eficientemente mediante sentencias del cálculo proposicional, de hallarse un método general y eficiente de resolver el problema de satisfactibilidad de las sentencias del cálculo se podrían resolver los problemas de clase *NP*. Tal método no ha sido hallado. La fórmula se satisface suponiendo que sólo sean verdaderos los enunciados. "*X* es rojo", "*Y* es azul", "*Z* es verde" y "*W* es rojo".



EL PROBLEMA DEL ARBOL pide hallar la mínima red de líneas que interconecte un conjunto de puntos, o lo que es equivalente, el mínimo sistema ferroviario que enlace entre sí un conjunto de ciudades *(a)*. Si los tramos de vía solamente pueden unirse en ciudades, el problema puede resolverse eficientemente por un procedimiento llamado algoritmo avaro. Primero se enlazan las dos ciudades más próximas; después, las siguentes en proximidad, y así sucesivamente *(b)*. Las ciudades ya conectadas indirectamente entre sí no vuelven a unirse (como en la línea *4* del ejemplo). El resultado es el llamado árbol óptimo *(c)*. Puede construirse una red todavía más corta si está permitido enlazar líneas en puntos aislados. En tales nudos se deben reunir tres líneas formando ángulos de 120 grados, pero no se conocen algoritmos eficientes para determinar dónde deben situarse los nudos de enlace. La red óptima, que aparece ilustrada abajo a la derecha, se obtuvo mediante exploración exhaustiva.

demostración que es demasiado larga para servir como certificado de pertenencia a la clase *NP*. Por otra parte, el complementario del problema del número compuesto, que pediría determinar si un número es primo, resulta pertenecer a la clase *NP*. La razón, que no es en modo alguno evidente, se debe a que en 1975, Vaughan Pratt, del Instituto de Tecnología de Massachusetts, descubrió métodos relativamente breves de probar el carácter primo de un número. A pesar de ello, se desconoce si el problema del número compuesto y su complementario pertenecen a la clase *P*.

Es fácil demostrar que todo problema de la clase *P* pertenece también a la clase *NP*. Si un problema pertenece a *P*, por definición existe para él un algoritmo eficiente. Para obtener una demostración breve y convincente de que la respuesta a un caso particular del problema es afirmativa, todo lo que se necesita es aplicar el algoritmo; el registro de los pasos del proceso constituye el certificado requerido.

La clase *NP* puede definirse también como la clase de problemas cuya respuesta es "sí" o "no", y que podrían resolverse brevemente por tanteo. Si se tiene un caso particular, de problema *NP* cuya respuesta resulte ser afirmativa, tal respuesta podría descubrirse bastante pronto realizando unos cuantos tanteos si la suerte ayudase un poco; si su respuesta fuese negativa, no se podría hallar por tanteo tal respuesta antes que si se efectuara una exploración exhaustiva. Por ejemplo, al estudiar un problema de Hamilton concreto, se podría encontrar un camino correcto (si es que hay alguno) al primer intento; trazando pequeñas porciones del itinerario y tanteando en cada etapa la forma de proseguir. Tal proceder, y vale la pena insistir en ello, no es algorítmico. Para convertirlo en un algoritmo sería necesario tachar cada camino ensayado después de verificarlo, y examinar todos los caminos posibles, método equivalente al de exploración exhaustiva.

Puede parecer grotesco definir un método matemático en términos de felices ideas y tanteos afortunados, pero, no obstante, es un método perfectamente legítimo de precisar cuáles son los problemas que pertenecen a la clase *NP*. En teoría, se podría incluso mecanizar el procedimiento construyendo un artificio denominado "máquina de Turing no determinística". Este dispositivo puede efectuar todas las funciones de una máquina de Turing ordinaria, pero además, en algunos puntos de su funcionamiento, puede elegir de más de una manera cuál va a ser la etapa siguiente. Se consideraría que tal máquina habría respondido afirmativamente a una cuestión si existiera alguna sucesión de elecciones que pudieran llevarla a conclusión afirmativa. La clase *NP* de los problemas no-determinísticamente polinómicos está formada precisamente por aquellos problemas cuyos casos particulares de solución afirmativa podrían ser identificados por máquinas efectuando una sucesión relativamente breve de cómputos de tanteo.

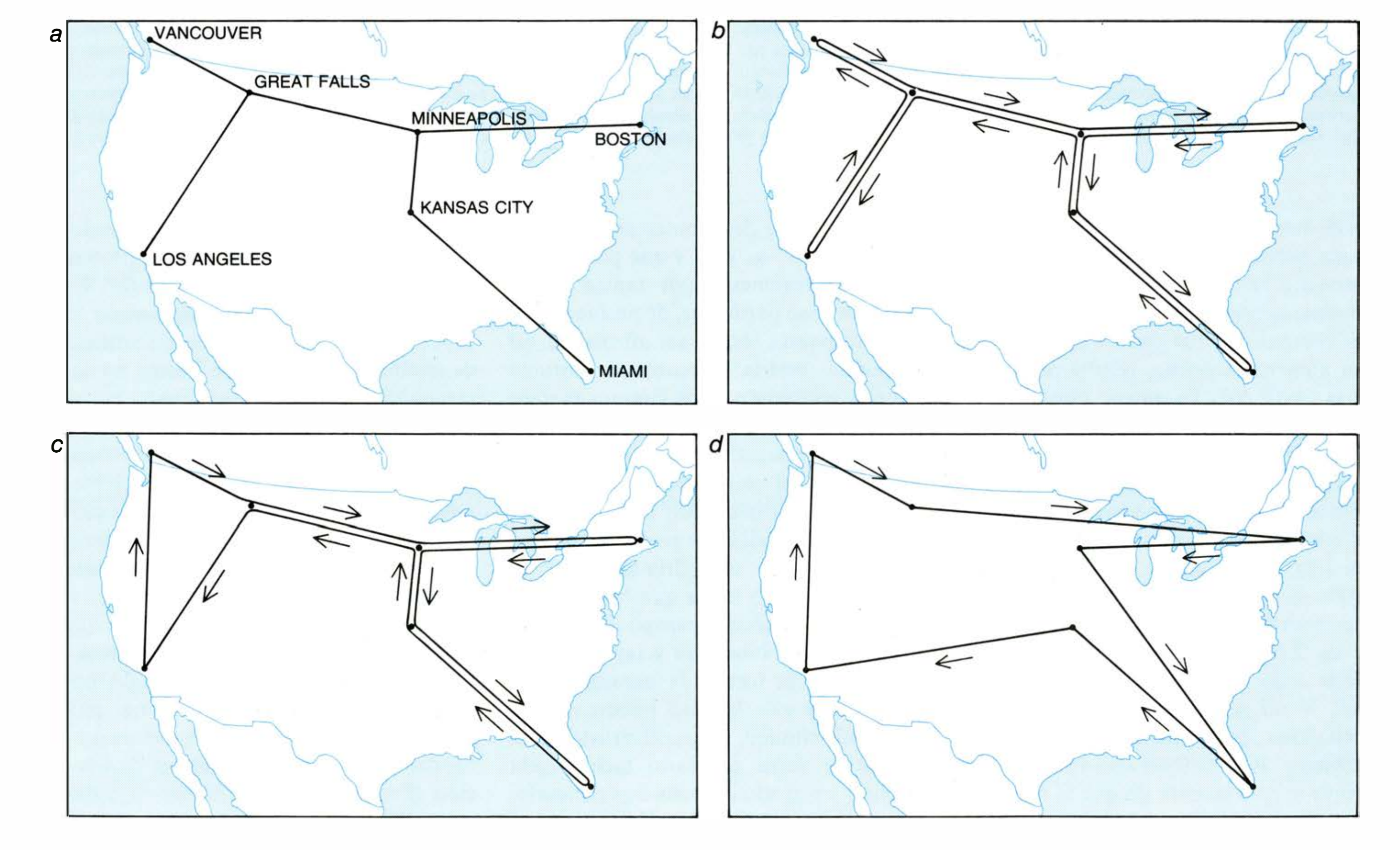
La posibilidad de definir problemas como los anteriores, apelando a la noción de tanteo afortunado, ha inducido

a muchos matemáticos a opinar que las clases *P* y *NP* no son el mismo conjunto, y que, por consiguiente, nunca podrán encontrarse algoritmos eficientes para los problemas no manejables de la clase *NP*. Si todo problema perteneciente a *NP* perteneciera en realidad a *P*, todo el trabajo de "tanteo y adivinación" podría sustituirse sin excesivo sacrificio de tiempo por un proceder sistemático. Cuesta trabajo que la capacidad de adivinar y la posibilidad de tener suerte puedan aportarnos tan poco.

La clase *NP* contiene una diversidad de problemas de frecuente aparición, y que parecen desafiar la posibilidad de resolverlos eficientemente. Hemos mencionado ya el problema de Hamilton y el problema de los números compuestos. Otro ejemplo es el llamado problema de emparejamiento. Podemos comprenderlo imaginando el problema que todos los meses de septiembre se plantea a los colegios mayores de las universidades, al tener que asignar habitaciones dobles a los estudiantes recién llegados.

Supongamos para mayor sencillez que toda la información referente a los hábitos de los recién llegados (si son fumadores o no, hábitos de sueño, gustos musicales, etcétera) pueda condensarse en una única decisión de compatibilidad o incompatibilidad de cada posible par de estudiantes. Todo el conjunto de nuevos estudiantes puede entonces representarse mediante un grafo, donde cada estudiante esté representado por un punto, y cada par de estudiantes que puedan compartir la misma habitación estén unidos por una línea. Si en cada habitación van a alojarse solamente dos estudiantes, el reparto puede realizarse eficientemente utilizando un ingenioso algoritmo de tiempo polinómico descubierto por Jack Edmonds, de la Universidad de Waterloo. No se conoce, sin embargo, ningún algoritmo eficiente para resolver este problema si cada habitación ha de ser compartida por tres estudiantes. El problema pertenece a la clase *NP*, pues en todos los casos de solución afirmativa se pueden dar certificaciones breves: una vez descubierto, es fácil exhibir el reparto de habitaciones. Evidentemente, las soluciones podrían hallarse por examen exhaustivo de todas las posibilidades, aunque este método sería muy poco eficiente. Si hay un poco de suerte, se podrá hallar pronto un reparto satisfactorio, si es que hay alguno.

El coloreado de mapas es un problema de la clase *NP* que preocupa a los matemáticos mucho más que a los cartógrafos. El problema consiste en determinar si se podrán colorear todos los países de un mapa dado usando un número especificado de colores distintos, de forma que ningún par de países fronterizos tengan el mismo color. Es sencillo decidir si un mapa podrá ser coloreado con dos colores: es necesario y suficiente que en ningún lugar del mapa se toquen en un punto un número impar de países. Todavía es más fácil decidir si un mapa puede colorearse con cuatro tintas. De hecho, ni siquiera es necesario ver el mapa, pues Kenneth Appel y Wolfgang Haken, de la Universidad de Illinois, demostraron, en 1975, que cuatro colores bastan para todo mapa. Sorprendentemente, sin embargo, no se dispone de ningún algoritmo eficiente que permita decidir si un mapa dado se podrá ilumi-



EL PROBLEMA DEL CIRCUITO MINIMO no puede ser resuelto por ningún algoritmo eficiente conocido. Los matemáticos han diseñado soluciones buenas, aunque no óptimas. Existe un procedimiento eficiente para trazar un itinerario que con certeza no sea mayor que el doble del mínimo. El método comienza hallando el árbol óptimo *(a)*, que puede determinarse eficientemente mediante el algoritmo avaro. El árbol se convierte en circuito recorriendo dos veces cada línea, una en cada sentido *(b)*. Este circuito mide exactamente el doble que el árbol mínimo; a su vez, el árbol mínimo ha de ser más corto que cualquier circuito de las ciudades, pues todo circuito puede convertirse en árbol (aunque un árbol sin ramas), suprimiendo un tramo entre dos ciudades. El circuito obtenido por este método puede, por lo general, mejorarse con ciertos atajos *(c, d)*.

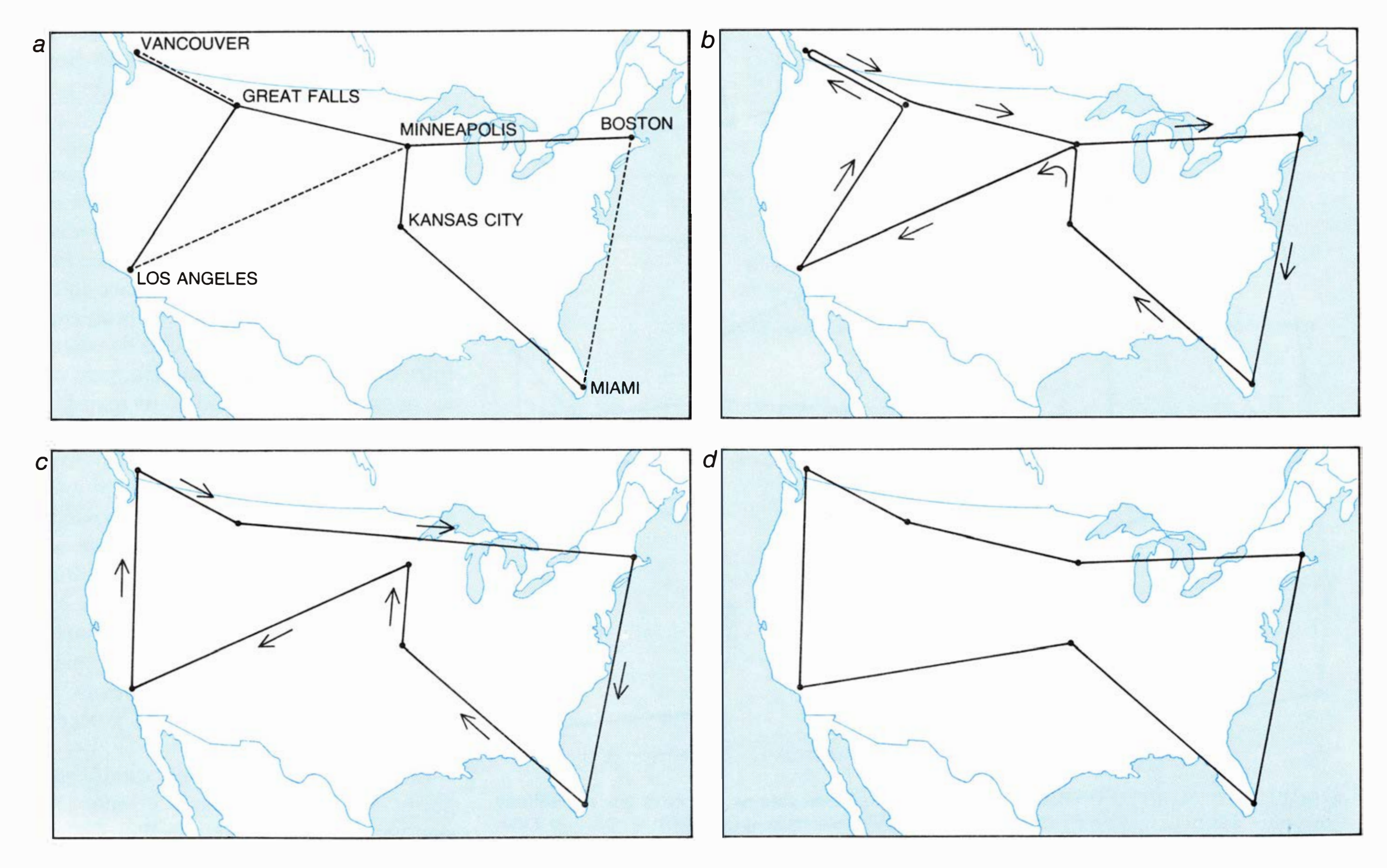
nar con tres colores. El problema pertenece a la clase *NP*, pues un correcto entintado del mapa serviría para certificar la respuesta afirmativa.

El coloreado de mapas puede considerarse un caso particular de otro problema llamado coloreado de grafos. Cualquier mapa puede convertirse en un grafo, reduciendo cada país a un punto y trazando una línea que una todo par de puntos cuyos correspondientes países tengan frontera común. Colorear el mapa es equivalente a la condición de colorear el grafo, si se impone la condición de que puntos unidos por una línea han de tener distinto color. El coloreado de grafos es, sin embargo, un problema más general que tiene aplicaciones fuera de la teoría de grafos. Por ejemplo, un grafo puede representar el plan de trabajo de una fábrica. Cada punto del grafo representa una cierta tarea a efectuar; dos puntos unidos por una línea representan tareas que no pueden ejecutarse simultáneamente, quizá, por ejemplo, porque necesitan la misma maquinaria. Un coloreado del grafo con tres colores proporcionaría un plan que dividiría el trabajo de la fábrica en tres turnos. Lo mismo que para mapas, el coloreado de grafos pertenece a la clase *NP*.

Ocurre con frecuencia que si un problema puede ser eficientemente resuelto, lo mismo es cierto para otros muchos. Por ejemplo, si se pudiera encontrar un algoritmo eficiente para el problema de colorear grafos, tal algoritmo podría aplicarse con ligeras modificaciones al problema del coloreado de mapas y de la planificación industrial. Así pues, se dice que el coloreado de mapas y la planificación industrial son eficientemente reducibles al coloreado de grafos. En estos últimos años ha quedado de manifiesto que ciertos problemas de la clase *NP* tienen esta notable propiedad: todos los problemas de *NP* son eficientemente reducibles a ellos. Estos problemas de "élite" en la clase *NP* se denominan *NP*-completos. Si alguno de ellos admite un algoritmo eficiente, entonces todo problema de *NP* puede ser eficientemente resuelto.

La primera demostración del carácter *NP*-completo de ciertos problemas fue presentada por Stephen A. Cook, de la Universidad de Toronto, en 1971. Su razonamiento discurre en línea esencialmente paralela a la de anteriores trabajos de Turing sobre las máquinas matemáticas y la relación de éstas con problemas de la lógica formal. Para dar su demostración Cook recurrió al cálculo proposicional, que es un lenguaje formal en el que enunciados lógicos individuales, cada uno de los cuales puede ser verdadero o falso, se ensamblan mediante las partículas conectivas "y", "o" y "no". En general, una sentencia del cálculo proposicional resulta ser verdadera o resulta ser falsa, dependiendo de cuáles de sus enunciados componentes se supongan verdaderos o falsos. Algunas sentencias, sin embargo, no pueden ser verdaderas en ninguna de las interpretaciones de sus componentes porque son contradictorias en sí mismas. Las sentencias que nunca pueden resultar verdaderas se llaman insatisfactibles, o lógicamente falsas.

Cook utilizó el cálculo proposicional para describir el funcionamiento de las máquinas de Turing no determinísticas, que son los dispositivos mecánicos de



ALGORITMO PERFECCIONADO para el problema del representante, que da un recorrido cuya longitud nunca supera a la óptima en más del 50 por ciento. El procedimiento fue preparado por Nicos Christofides, del Imperial College of Science and Technology. El primer paso vuelve a ser la obtención del árbol mínimo. Todas las ciudades enlazadas por el árbol con un número impar de ciudades se singularizan; en este ejemplo todas las ciudades, a excepción de Kansas City, tienen un número impar de enlaces. Estas ciudades distinguidas se conectan por pares de forma parecida a la de emparejar estudiantes, con lo que se obtiene un circuito *(a)* que se mejora tomando atajos *(b, c)*. El resultado es un poco más largo que el óptimo *(d)*, hallado por exploración exhaustiva. En el circuito mínimo no figura el enlace entre las dos ciudades más cercanas.

tanteo, esenciales para definir la clase *NP.* Cook demostró que los cálculos que pudiera realizar cualquiera de estas máquinas podrían ser descritos sucintamente por fórmulas del cálculo proposicional. Cuando a la máquina se le presenta un caso particular con respuesta afirmativa de un problema *NP,* su funcionamiento puede describirse por una fórmula satisfactible. En cambio, si el caso particular anterior tuviera respuesta negativa, la fórmula proposicional resultante sería no satisfactible.

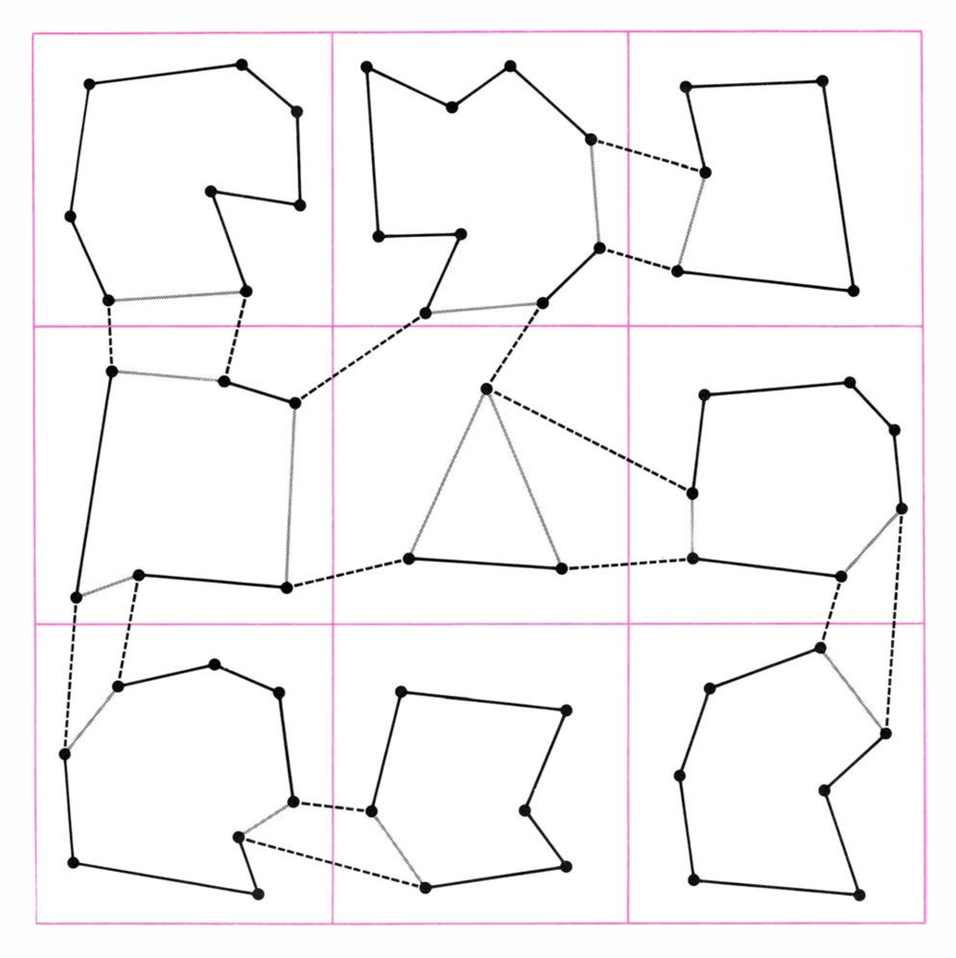
De la demostración de Cook se desprende que si se pudiera determinar eficientemente si una fórmula del cálculo proposicional es satisfactible, se podría también determinar eficientemente, y por adelantado, si un problema presentado a una máquina de Turing no determinística tendrá respuesta afirmativa o negativa. Como los problemas de la clase *NP* son precisamente, por definición, todos los resolubles mediante máquinas de Turing no determinísticas, se dispondría de un método eficiente de resolver todos los problemas de esta clase. Evidentemente, el fallo está en que no se conoce ningún método eficiente para decidir si una fórmula del cálculo proposicional es satisfactible o no.

En esencia, el razonamiento de Cook establece que el cálculo proposicional constituye un lenguaje universal para describir los problemas de la clase *NP.* Todo caso particular de uno de tales problemas se corresponde con una fórmula de dicho lenguaje y, si la fórmula es satisfactible, el problema tiene respuesta afirmativa. Se ha demostrado desde entonces el carácter *NP*-completo de otros muchos problemas, lo que se ha conseguido demostrando que el problema de determinar si una fórmula del cálculo proposicional es satisfactible puede reducirse eficientemente a ellos. El problema de Hamilton, el problema de acoplar grupos de tres compañeros de habitación, y el problema de colorear grafos con tres tintas son todos ellos *NP*-completos. El primero en señalar el amplio campo de aplicación de esta teoría fue Richard M. Karp, de la Universidad de California en Berkeley. Se han realizado también otras investigaciones parecidas bajo la dirección del matemático ruso P. A. Levin. Dado que los problemas *NP*-completos encierran en sí todas las dificultades de los demás problemas de *NP,* está hoy ampliamente difundida la convicción de que los problemas *NP*-completos son computacionalmente imposibles de tratar. Por lo general, si llega a demostrarse que un problema es *NP*-completo, se acepta tal razón como suficiente para abandonar la búsqueda de algoritmos eficientes que los resuelvan.

Ni aún admitiendo que ningún problema *NP*-completo puede ser computacionalmente manejable quedarían zanjadas todas las cuestiones relativas a la clase *NP.* Hay zonas aún más obscuras que los misteriosos problemas *NP*-completos: problemas de la clase *NP* para los que no se conocen algoritmos eficientes ni tampoco ha podido demostrarse que sean *NP*-completos. El problema de los números compuestos es uno de ellos.

No todos los problemas resolubles por computador tienen respuestas del tipo "sí" o "no". Otro tipo común de problemas son los de optimización. Por ejemplo, supongamos que se nos den las situaciones en el mapa de ciertas ciudades, y que se nos pida determinar la red mínima de ferrocarriles que las enlace entre sí. En una de las versiones del problema se permite trazar líneas rectas de enlace entre dos ciudades cualesquiera, pero no se pueden instalar nudos ferroviarios aislados: los tramos de vía han de unirse necesariamente en ciudades. Hay una propiedad de la solución de este problema que es obvia: la red óptima no puede contener circuitos cerrados, porque, de hacerlo, se podría acortar la longitud total de la red suprimiendo uno de los tramos de enlace en el circuito. Así pues, la red óptima siempre se ramifica, como un árbol, por lo que los problemas de este tipo se llaman problemas arborescentes.

El problema del árbol ramificado puede resolverse correcta y muy eficientemente por un método llamado algoritmo avaro, diseñado por Joseph B. Kruskal, de los Bell Laboratories. El procedimiento consiste, simplemente, en conectar entre sí primero el par de ciudades más próximas una de otra, des-



LA SOLUCION APROXIMADA del problema del representante no garantiza que el resultado se encuentre dentro de un margen especificado *a priori* del recorrido óptimo, pero al aplicarlo a muchos casos particulares del problema pocas veces da soluciones muy alejadas de la óptima. El algoritmo, diseñado por Richard M. Karp, de la Universidad de California en Berkeley, descompone el mapa en pequeñas regiones, cada una de las cuales contiene solamente unas cuantas ciudades. Dentro de cada región se encuentra el circuito mínimo por exploración exhaustiva, procedimiento que ahora es viable porque el número de ciudades es suficientemente pequeño. Cada uno de los pequeños recorridos se considera como entidad única a enlazar con las otras, tarea que puede realizarse por un algoritmo eficiente, parecido al algoritmo avaro para hallar el árbol óptimo.

pués las siguientes en proximidad, y así sucesivamente, evitando añadir líneas superfluas (líneas que enlazan ciudades ya conectadas indirectamente entre sí). No es en modo alguno evidente que este método proporcione siempre en todos los casos la red de longitud mínima, pero el caso es que lo hace, y tiene además la agradale propiedad de no exigir, ni prever, ni reconsiderar decisiones anteriores.

Podemos fiarnos del algoritmo avaro para determinar la red mínima entre ciudades con las condiciones anteriormente especificadas, pero, en general, la red así obtenida no será la mínima posible. Se puede economizar todavía más introduciendo nudos de enlace entre las ciudades. Las propiedades de las redes con nudos de enlace intermedios fueron estudiadas por el matemático suizo Jakob Steiner. Puede demostrarse que en toda red mínima es preciso disponer los nudos de enlace de forma que se compongan de tres líneas que se encuentren a 120 grados. Esta regla sirve de orientación para evaluar posibles redes, pero hay muchas redes posibles con nudos de enlace que cumplen la condición de Steiner. No se ha descubierto ningún algoritmo que encuentre rápidamente la red óptima.

El problema del itinerario del viajante está íntimamente relacionado con el anterior. Igual que antes, se tiene una colección de ciudades, pero ahora se pide hallar el circuito de longitud mínima. Un primer intento, inspirado en el algoritmo avaro, sería hacer que el viajante se dirigiese en cada caso a la ciudad más próxima de las que todavía no haya visitado, pero este procedimiento no resuelve el problema. De hecho, el problema es notable porque ha resistido todos los intentos de resolverlo eficientemente.

Los problemas de optimización no son sustancialmente diferentes de los problemas que plantean disyuntivas a resolver afirmativa o negativamente. En el problema del viajante, por ejemplo, se nos podría dar la longitud de un circuito que pasa por varias ciudades, y preguntársenos si puede, o no, construirse otro cuya longitud sea menor que la dada. La elucidación de esta disyuntiva no puede ser más difícil que su problema de optimización asociado, porque si pudiera hallarse un circuito mínimo por algún método eficiente, sería trivial decidir si su longitud es menor, o no, que un número dado. Por consiguiente, si la versión "disyuntiva" del problema es computacionalmente intratable, tampoco se puede resolver eficientemente el problema de optimización. Por esta razón se dice que ciertos problemas de optimización, como el del itinerario del representante o el emplazamiento de los puntos de enlace de Steiner, son *NP*-completos.

Se encuentran frecuentemente problemas de optimización en ingeniería, economía, investigación operativa y otros campos. Por consiguiente, el descubrimiento de que algunos de estos problemas son *NP*-completos reviste considerable interés práctico. Como verosímilmente los problemas *NP*-completos carecen de algoritmos eficientes, no parece tener ningún objeto invertir más tiempo y esfuerzo en la búsqueda de soluciones óptimas. La alternativa que se ha adoptado recientemente consiste en buscar soluciones aproximadas que sean buenas aún cuando no sean precisamente óptimas.

Se ha aplicado al problema del viajante una técnica que, si bien no da la solución óptima, sí garantiza una solución que nunca mide más del doble que la solución óptima. El proceso comienza construyendo el árbol ramificado mínimo, que puede conseguirse eficientemente con el algoritmo avaro. La red así obtenida puede convertirse en un recorrido por todas las ciudades atravesando cada línea dos veces y regresando al origen. Se sabe que el árbol ramificado mínimo ha de ser más corto que cualquier posible circuito que pase por todas las ciudades, pues un circuito puede convertirse en un árbol (aunque sin ramas) suprimiendo uno de los tramos del recorrido. Así pues, el doble de la longitud del árbol ramificado mínimo no puede ser más del doble del circuito mínimo. El método se ejecuta mediante un algoritmo de tiempo polinómico. Recientemente, Nicos Christofides, del Imperial College of Science and Technology de Londres, ha descubierto un método para mejorar el algoritmo, de forma que el circuito que produce nunca sea más de vez y media el óptimo.

Sacrificando la condición de que las soluciones sean óptimas, y renunciando a insistir en que las soluciones menos que óptimas se encuentren siempre dentro de un entorno especificado de aquéllas, pueden obtenerse soluciones aproximadas en un sentido radicalmente distinto. A cambio de tales sacrificios puede garantizarse que muy pocas veces las soluciones se alejarán mucho de la solución óptima.

La hipótesis subyacente a estas técnicas es que los mapas que puedan encontrarse en la realidad no están preparados con la intención específica de invalidar ideas fundamentalmente razonables; tales mapas "patológicos" sólo aparecen con suficiente frecuencia cuando son específicamente diseñados por los especialistas de ciencias de la computación con la intención de mostrar los fallos del trabajo de sus colegas. En realidad, si los algoritmos antes comentados para el problema del representante se aplican a mapas "naturales", dan de sí mucho más de lo que prometen. Los circuitos resultantes no son el 100 o el 50, sino solamente alrededor de un 5 por ciento más largo que los óptimos.

Para la mayor parte de los mapas puede suponerse razonablemente que las ciudades están en ellos situadas al azar. En 1958, Jillian Beardwood, J. H. Halton y John M. Hammersley, de la Universidad de Oxford, demostraron un teorema que describe las propiedades estadísticas de los circuitos óptimos que pasan por puntos aleatoriamente distribuidos. Apoyándose en ese teorema, Karp ha demostrado que hay un método sencillo de construir recorridos que da casi siempre resultados óptimos al aplicarlo a mapas con muchas ciudades.

Karp comienza descomponiendo el mapa en muchas regiones pequeñas. Dentro de cada una de estas regiones hay un número lo bastante pequeño de ciudades como para poder determinar el circuito óptimo por exploración exhaustiva, a pesar de que este método utiliza un algoritmo de tiempo exponencial. Los recorridos de las regiones pequeñas se interconectan mediante una variante del algoritmo avaro. Quizá significativamente, el método no es muy diferente del empleado para resolver manualmente este problema.

Se pueden obtener soluciones eficientes, aunque aproximadas solamente, de muchos problemas de optimización *NP*-completos. Sin embargo, desde el punto de vista de las matemáticas, la cuestión importante es si *NP* es una clase idéntica a *P*. Los reiterados fracasos que han sufrido todos los intentos de hallar algoritmos eficientes para los problemas *NP*-completos han hecho crecer la confianza en que *NP* y *P* no sean clases iguales. Se tiene hoy la sospecha de que no son idénticas, pero es posible que la demostración de su distinto carácter no esté al alcance de la matemática actual. Esta cuestión puede muy bien sumarse al selecto grupo de enigmas matemáticos que permanecen sin solución durante varias décadas, en espera del desarrollo de nuevos métodos.



UN AURIGA DE LOS AZULES, con el látigo sostenido en alto por su mano derecha y las riendas sujetas con la izquierda, retratado en mosaico. Es uno de los cuatro aurigas que aparecen en las puertas de salida de un circo. El mosaico decoraba el umbral de un comedor excavado recientemente en el solar de una casa de la Cartago romana, donde ha comenzado una campaña internacional de arqueología de urgencia.

La Cartago romana

Una campaña internacional de arqueología de urgencia en las ruinas de la ciudad está descubriendo importante información sobre sus sucesivos papeles como colonia romana, botín de los vándalos y avanzada del Imperio Bizantino

John H. Humphrey y John Griffiths Pedley

En el último cuarto del primer milenio a. C., Cartago, la mayor de las colonias fenicias en Africa, y Roma, entonces joven república expansionista de la Italia central, lucharon encarnizadamente por el dominio del Mediterráneo occidental. Entre 264 y 146 a.C. las dos potencias guerrearon entre sí en tres ocasiones. Aunque unos 50 años de paz inestable separaron el fin de la Segunda Guerra Púnica y el comienzo de la Tercera, el Senado romano no pudo olvidar la rivalidad. Los senadores oían a Catón el Mayor exigir en toda posible ocasión que Cartago debía ser destruida.

Catón no vivió lo bastante para ver cumplida su voluntad; murió a la edad de 85 años en el 149 a.C., en los instantes en que Roma preparaba su campaña definitiva contra los cartagineses. Tras un penoso asedio de tres años, Escipión Emiliano, nieto adoptivo del jefe romano vencedor de la Segunda Guerra Púnica, conquistó Cartago en el 146 a.C. Sus primeras órdenes, una vez que los supervivientes de la ciudad fueron dispersados, hubieran merecido los plácemes de Catón. Las tropas de Escipión arrasaron todos los edificios de la ciudad; el lugar fue entonces maldecido y sembrado ritualmente de sal.

El aparente carácter definitivo del gesto de Escipión fue, sin embargo, efímero. En menos de 25 años los romanos dieron la vuelta a su política. Reconstruida como colonia romana, Cartago volvió a ser una vez más una gran ciudad, destinada a sobrevivir otros 800 años. La brillante historia de estos siglos púnicos y postpúnicos está siendo hoy leída en las ruinas de la ciudad como resultado de una empresa internacional sin precedentes en la arqueología de urgencia. Patrocinado por el director general de la UNESCO (United Nations Educational, Scientific and Cultural Organization) en 1972, el proyecto incluye a estudiosos no sólo del país anfitrión, Túnez, sino también de otras nueve naciones de Europa y América.

Un breve bosquejo de la historia de la Cartago romana basta para demostrar que hay tarea más que suficiente para todos estos participantes. Como capital de la provincia romana de Africa a finales del siglo primero a.C., la nueva ciudad de Cartago llegó a convertirse en la tercera metrópoli más rica del Imperio Romano, rivalizando con la misma Roma en magnificencia. El emperador Augusto rebautizó a la ciudad como Colonia Julia Carthago, y con el advenimiento del cristianismo se convirtió en un centro de fermentación religiosa. En el 439 d.C., año en que Cartago se rindió a los vándalos invasores, la ciudad tenía por lo menos 22 grandes iglesias, su propio obispo, más de otros 500 clérigos y una población que se calcula alrededor de 250.000 almas, o incluso más.

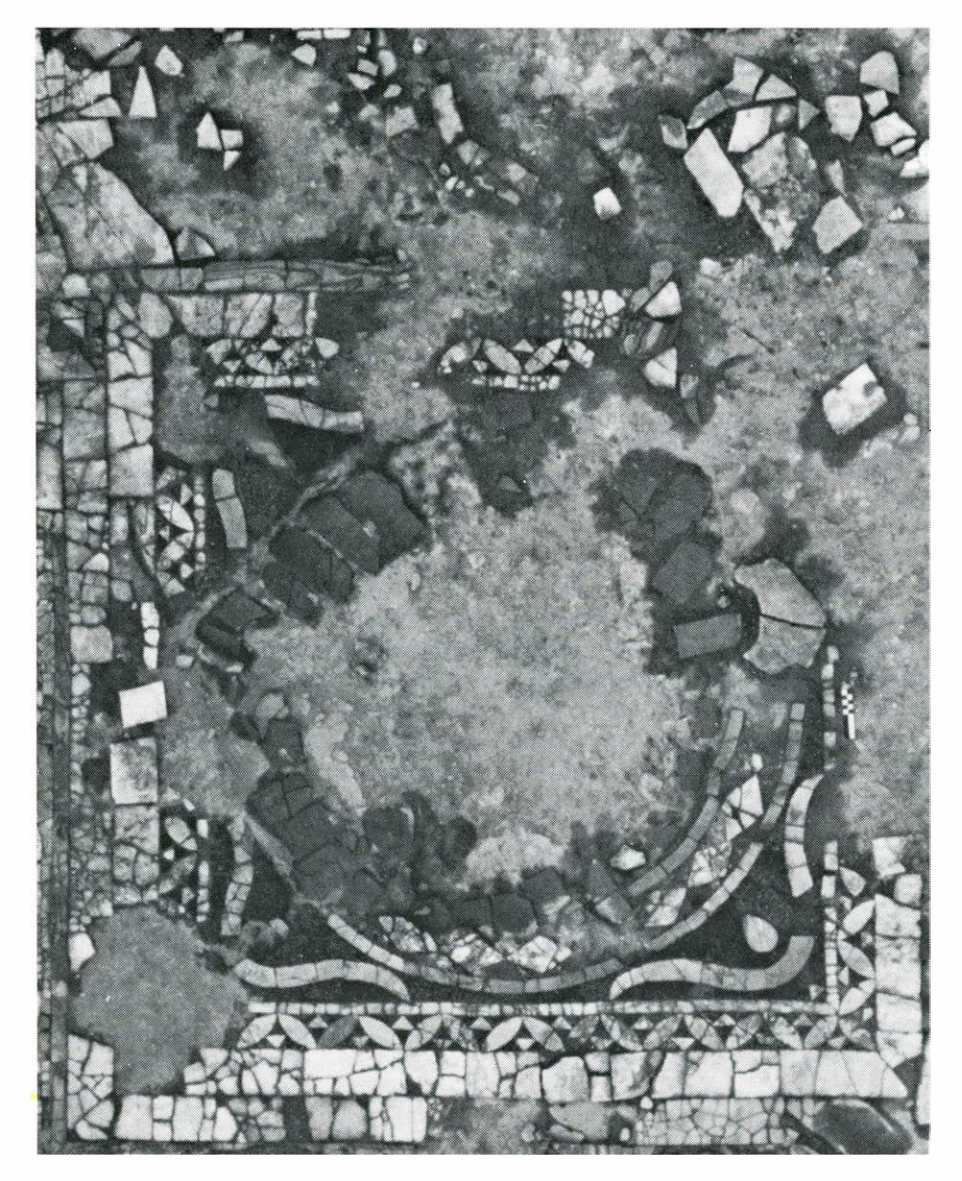
Las tropas imperiales acabaron por expulsar a los vándalos, pero eran tropas del Imperio Romano de Oriente; la ciudad liberada recibió el nuevo nombre de Colonia Justiniana Carthago en honor del monarca de Bizancio, Justiniano el Grande. Cartago siguió siendo una ciudad bizantina desde 533 d.C. hasta 697. Pudiera haber subsistido hasta el día de hoy si la población no se hubiese sublevado en 697 contra una pequeña guarnición árabe puesta allí por el gobernador de Egipto, Hassan ibn en Noman. Según refiere el estudioso clásico francés E. C. F. Babelon, Hassan, cuyas fuerzas se ocupaban de incorporar Argelia y Marruecos a los crecientes dominios del Islam, montó en cólera por lo que él consideraba un ataque a traición. Repitiendo en el año 698 d.C. la devastación romana de cerca de 1000 años antes, logró que la ciudad de Cartago quedase una vez más en ruinas.

Hoy en día, unos 1300 años más tarde, la ciudad destruida por dos veces se ha convertido en un moderno solar de edificación. Sus hectáreas vacías van desapareciendo bajo la expansión suburbana de la próxima Túnez. En 1881, Túnez fue ocupada por los franceses y, en los años siguientes a la Segunda Guerra Mundial, lo que había sido un goteo de desarrollo urbano se convirtió en una inundación. A comienzos de la década de 1970 la abandonada Cartago era el lugar de urbanización más codiciado de todo el país de Túnez. El nuevo palacio del presidente se había construido sobre las ruinas del norte de Cartago y grandes chalets aparecían en las alturas que dominan la costa donde se había levantado la antigua Byrsa, como se llamaba la ciudadela de la ciudad púnica. Las lomas adyacentes, conocidas en época romana como la Colina de Juno y donde todavía se aprecian las ruinas de un teatro romano, se vieron también invadidas. Otros varios chalets se edificaron a lo largo de la costa al norte de Túnez hasta más allá de la antigua muralla de la ciudad romana, creando los actuales suburbios de Salambó, Dermech y Sayda.

Cuando el programa de arqueología de urgencia fue puesto en práctica a mediados de la década de 1970, los diversos sectores no urbanizados de la ciudad en ruinas fueron asignados a los varios arqueólogos participantes. El Instituto Nacional de Arqueología y Arte de Túnez, bajo la dirección de Azedine Beschaouch, y Abdelmajid Eunatti, conservador del yacimiento de Cartago, han llevado adelante trabajos anteriores en las villas romanas cerca del teatro romano. Sin embargo, la empresa principal del país anfitrión ha sido un vasto programa de catalogación que abarca no sólo los descubrimientos actuales, sino también los hallazgos esporádicos de los últimos 150 años. A un grupo británico, dirigido por Henry Hurst, le fue asignada la isla situada en el centro de

uno de los dos puertos artificiales de la ciudad: un puerto circular considerado tradicionalmente como militar en sus tiempos. Los británicos han recibido también un segundo yacimiento al lado norte de aquel puerto, y una tercera zona a corta distancia tierra adentro, sobre la línea de la muralla de la ciudad romana.

Desde el puerto circular hacia el norte y hasta los límites de la ciudad, sólo quedaba una pequeña zona sin ocupar en la costa del Golfo de Túnez (excepto las ruinas de las Termas Antoninianas, parte de la Cartago romana que fue convertida en parque arqueológico hace unos años); ésta fue asignada a excavadores alemanes bajo la dirección de F. Rakob. Otros varios grupos se concentraron en partes de la ciudad situadas a mayor distancia del mar: los franceses, a las órdenes de S. Lancel, prefirieron continuar trabajos anteriores de una serie de investigadores franceses en la ciudadela púnica; un grupo canadiense, dirigido por P. Sénay, se encargó de ampliar el trabajo realizado anteriormente por los franceses en un monumento circular de la Colina de Juno, y otro, bajo C. Wells y E. Wightman, siguió la muralla de la ciudad en sus proximidades. Aun más tierra adentro, un grupo italiano, a las órdenes de A. Carandini, se encargó de explorar la red de calles y la muralla urbana en el ángulo noroeste de la ciudad. Arqueólogos del Instituto de Historia de la Cultura Material, de Varsovia, hicieron el levantamiento del sector suroeste, particularmente en las cercanías del circo romano, midiendo las variaciones de la resistencia eléctrica del suelo. El grupo danés, comandado por S. Dietz, eligió un yacimiento costero al norte de las Termas Antoninianas, fuera de los límites de la ciudad, donde anteriormente se habían efectuado excavaciones danesas. Un grupo búlgaro trabajó en otro yacimiento de las afueras de la ciudad: Damous el Karita, una famosa iglesia paleocristiana.



PAVIMENTO DE MARMOL que cubría el suelo del comedor después de la amplia reconstrucción de la casa que siguió a la derrota de los vándalos por los bizantinos en 533 d.C. El pavimento, de la clase *opus sectile,* fue reparado dos veces por los ocupantes de la casa en el siglo siguiente.

Dos grupos americanos han intervenido en la excavación de urgencia. Uno de ellos, una empresa mixta de la Universidad de Harvard y de la Universidad de Chicago, recibió una zona en el borde occidental del segundo puerto de Cartago, un puerto rectangular que se ha venido considerando tradicionalmente como mercantil, situado inmediatamente al sur del circular. La zona se prolonga hacia el oeste por el *tophet* púnico, o lugar de enterramiento, y el santuario de la diosa cartaginesa Tanit. Nuestro propio grupo, de la Universidad de Michigan, se hizo cargo de un campo rectangular de 100 por 400 metros. La expansión urbana moderna había sido prevista en esta zona paralelamente a la vía férrea de Túnez a La Marsa, situada al oeste del puerto circular. Estábamos literalmente continuando una tradición de Michigan, pues el primer arqueólogo americano que trabajó en Cartago fue un arqueólogo clásico de la Universidad, Francis W. Kelsey; su nombre ha sido dado al museo universitario que nosotros representábamos como director de campo y principal investigador, respectivamente. El trabajo de los dos grupos americanos está patrocinado por las Escuelas Americanas de Investigación Oriental y financiado en gran parte por el programa de Moneda Extranjera de la Smithsonian Institution.

El padre, o quizá más propiamente el abuelo, de los modernos estudios cartagineses fue un erudito del siglo XIX, C. T. Falbe, que actuó como cónsul danés en Túnez y publicó en 1833 el primer plano detallado de las ruinas de la ciudad. Algunos otros aficionados a las antigüedades, incluido el capellán anglicano de Túnez, Nathan Davis, trabajaron en el yacimiento en los años siguientes; y poco después de la ocupación francesa, eruditos de esta nacionalidad elaboraron un proyecto global, pero disparatadamente ambicioso, para excavar todos los lugares más importantes de la ciudad.

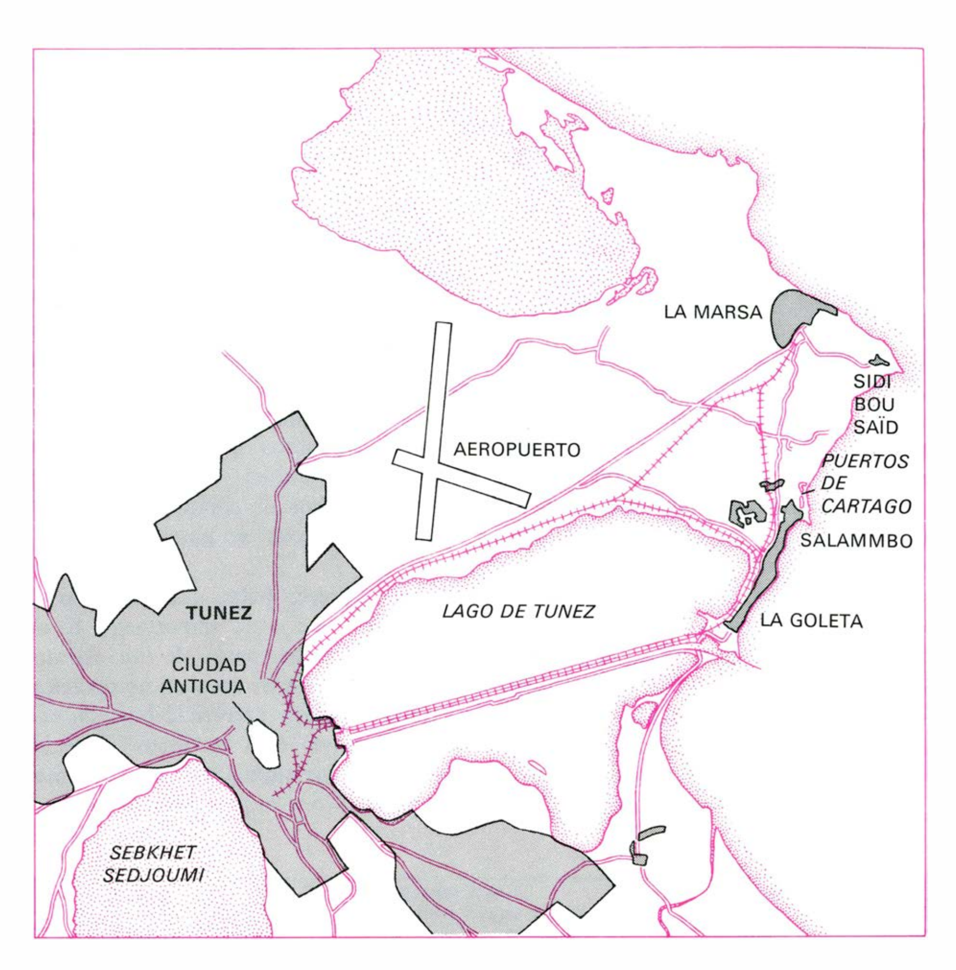
Los esfuerzos de estos investigadores y de otros que les sucedieron, sin embargo, sirvieron principalmente para llenar los museos de Europa de obras de arte tales como mosaicos y estatuas, cuyo exacto emplazamiento y contexto arqueológico quedaban generalmente sin documentar. En verdad, hasta que la presente campaña dio comienzo, la his-

toria registrada de excavación y documentación le dio fama de ser quizás el yacimiento clásico mediterráneo más difícil de estudiar. Aun el más terco estudioso de la Antigüedad que tratase de interpretar la obra realizada allí, se veía defraudado por la ausencia de datos estratigráficos, de análisis de cerámica, de estudios numismáticos e incluso de un instrumento tan fundamental como un mapa preciso de la ciudad a gran escala.

Las cosas están cambiando ahora. Cada uno de los grupos que trabaja en Cartago tiene sus propios intereses y especialidades particulares, que por primera vez están contribuyendo a una visión de conjunto de la larga historia de la ciudad. Por ejemplo, un grupo ha decidido concentrarse en su historia más reciente, desde el 400 d.C. aproximadamente, unos 40 años antes de la ocupación vándala, hasta 698 d.C. y los años siguientes, un período que abarca el resurgir bizantino de la ciudad como Colonia Justiniana y la captura final de la misma por los musulmanes.

El extenso campo concedido a nosotros era ideal para este propósito. El estudio preliminar del yacimiento por el Instituto Nacional de Arqueología y Arte de Túnez, realizado en 1971, mostraba que el campo incluía las ruinas de una iglesia cristiana tardía, con su baptisterio, y las de una serie de casas particulares ricas del período romano tardío (o sea, de Colonia Julia) y de las épocas siguientes vándala y bizantina. Ensanchando el enfoque tradicional de la arqueología clásica, conocido desde hace tiempo por su inclinación hacia instrumentos de la historia tales como inscripciones, monedas y productos de artistas y artesanos, nosotros contamos en nuestro grupo con especialistas interesados en la documentación relativa a la economía de la ciudad y al modo de vida de sus gentes. Por ejemplo, la alimentación de los cartagineses se refleja en los restos de plantas y animales, y su explotación del entorno físico aparece indicado en sus preferencias de piedra para construir y de arcilla para su cerámica.

El encuadre cronológico es preceptivo en el estudio de cualquier yacimiento clásico; permite a los investigadores relacionar los varios niveles de excavación y los objetos hallados en cada nivel con períodos conocidos de la historia. Cuanto mayor sea la precisión cronológica, más detalle podrá aplicarse a la interpretación de las diversas clases de datos. En Cartago, como en la mayoría de los yacimientos clásicos, las dos clases de objetos que proporcionan la base de un encuadre cronológico son las monedas y la cerámica. Por ejemplo, en razón de los intervalos de tiempo relativamente cortos que están implicados, las monedas suelen proporcionar un medio de fechar estratos que resulta mucho más preciso que pudieran serlo las determinaciones del carbono-14. Ahora bien: las monedas no se encuentran necesariamente en abundancia, mientras que la cerámica rota es probable que sea uno de los mayores componentes de los desperdicios de cualquier yacimiento clásico.



LAS RUINAS DE CARTAGO están situadas al norte de la moderna Túnez. En las últimas décadas, la expansión de los suburbios de Túnez en dirección norte casi ha sepultado la ciudad antigua.

En nuestra parcela oblonga éste resultó ser el caso una vez más. Las monedas que excavamos en 1975 y 1976 fueron identificadas rápidamente para nosotros por un colega de la Universidad de Michigan, T. V. Buttrey, y nos proporcionaron un marco cronológico de conjunto. Sin embargo, los tiestos cerámicos encontrados en relación estratigráfica con las monedas, y los otros miles de fragmentos de los diversos estratos intermedios del yacimiento, fueron los que nos permitieron efectuar una serie de subdivisiones cronológicas basadas en tipos y secuencias cerámicas.

El trabajo no carecía de dificultades. Hasta que la operación de rescate comenzó en Cartago, no había aparecido ni una publicación que describiese la estratigrafía relativa de la cerámica en parte alguna de la ciudad. Por ejemplo, la que se conoce como cerámica africana de barniz rojo, una clase fina producida en Cartago y sus alrededores, así como en otras localidades del norte de Túnez, era el principal servicio de mesa usado no sólo en Cartago sino también a través de todo el Imperio Romano durante medio milenio a partir del 100 d.C., aproximadamente. La historia de esta famosa cerámica fue laboriosamente reconstruida por John W. Hayes, del Museo Real de Ontario, en su libro *Late Roman Pottery.* Sin embargo, su obra definitiva hubo de basarse exclusivamente en los resultados de excavaciones realizadas fuera de Túnez, y así la clasificación de la cerámica se fundó, en su mayor parte, más en consideraciones estilísticas que cronológicas. En lo que a las monedas se refiere, hasta que empezaron las excavaciones de urgencia nunca se había publicado un catálogo de las monedas recuperadas en el suelo de Cartago.

Ha sido una suerte que Hayes y su co-

lega J. A. Riley, de la Universidad de Manchester, hayan tenido a su cargo el análisis de la cerámica de nuestro yacimiento. Como la fina cerámica de mesa de barniz rojo se reconoce con facilidad, nos ha proporcionado un marco general con el que relacionar las otras cerámicas finas y bastas. Los dos especialistas en cerámica hubieran podido fácilmente encontrarse enterrados en fragmentos de no haber seguido un plan de selección y eliminación. La tierra de cada "locus" (esto es, una determinada unidad del terreno) del yacimiento fue tamizada después de su extracción, y todos los fragmentos cerámicos recogidos en cubos para su lavado (y para su tratamiento con ácido cuando estaban muy impregnados de tierra). Los fragmentos de cada cubo fueron después clasificados según categorías preestablecidas. Algunos lo fueron en términos de su función, por ejemplo, trozos de candiles, jarros, fuentes, cuencos, ánforas y cacharros de cocina. Otras categorías se fundaban en el estilo de la cerámica, por ejemplo, cerámica de barniz mate, cerámica pintada y cerámica de barniz rojo.

El paso siguiente consistía en dividir los fragmentos de cada categoría en cuatro grupos: fragmentos de bordes, fragmentos de bases, fragmentos de asas y fragmentos de cuerpos de los recipientes. Los primeros cálculos cuantitativos se hacían en este momento: los fragmentos de cada uno de los cuatro grupos eran contados y pesados, y sus datos registrados. Si al procesarlos hasta este punto no había señales de que el yacimiento estaba alterado, Hayes y Riley consideraban seguro combinar todos los cubos de cerámica procedentes de cada "locus" en particular.

A continuación había que decidir si el "locus" era o no importante. Si el "locus" formaba parte de un estrato "cerrado" (es decir, un nivel no revuelto y completamente enterrado bajo un estrato superior) o si el "locus" parecía importante por alguna otra razón, como podría serlo por la presencia de una cerámica desusada, Hayes y Riley guardaban y registraban toda la cerámica recuperada en el mismo y asignaban números de catálogo a piezas selectas, que eran también medidas y fotografiadas o dibujadas. Si el "locus" ni era cerrado ni importante por otra razón, la mayoría de los fragmentos de cuerpos era desechada.

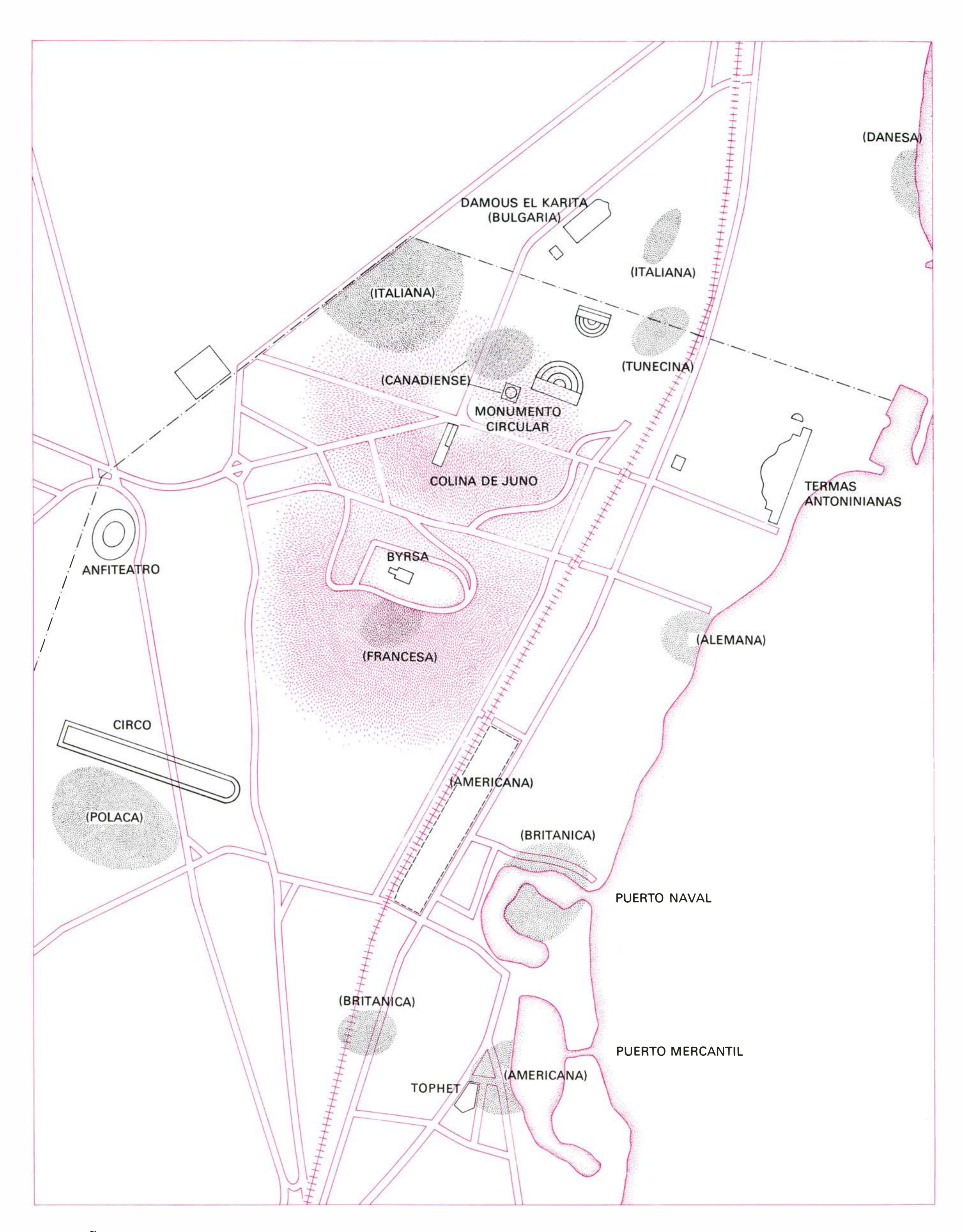
Como por regla general los fragmentos de cuerpos son mucho más numerosos que los de bordes, bases y asas, la cantidad real de cerámica que al final quedaba almacenada para su ulterior estudio era sólo una fracción de la cantidad excavada. Por ejemplo, a base de las monedas asociadas pudimos distinguir 81 variedades de cerámica en estratos pertenecientes al tercer cuarto del siglo sexto, un intervalo inicial del período bizantino de Cartago. De un total de 7343 fragmentos excavados, sólo 881 eran fragmentos de bordes, bases y asas, de manera que pudieron ser desechados más de 6000 fragmentos. La moneda más reciente asociada con este grupo de tiestos se acuñó en algún momento entre 548 d.C. y 565.

En 1976, Hayes y Riley publicaron los datos correspondientes a 17 grupos distintos o depósitos de cerámica de Cartago; las fechas de esta cerámica oscilaban desde principios del siglo primero hasta la segunda mitad del sexto. El año pasado siguieron publicando varios grupos cerámicos adicionales, la mayor parte de ellos pertenecientes al siglo séptimo.

Entre sus descubrimientos se constata un descenso en la importación de cerámicas finas de Italia, apenas iniciada la historia de Colonia Julia. Durante el siglo primero, la cerámica fina aretina que se importaba ascendía al 27 por ciento de tales cerámicas en nuestro yacimiento, y las ánforas importadas sumaban otro 2,5 por ciento. Esa proporción de cerámica importada descendió a una fracción de aquel porcentaje a comienzos del siglo segundo.

El procedimiento de tamizar la tierra nos dio mucho más que sólo tiestos. Nos permitió recuperar un número mucho mayor de monedas de las que pudieran haber aparecido usando métodos de excavación más tradicionales. Por ejemplo, recobramos más de 600 monedas en nuestra campaña de 1975 y más de 1200 en la de 1976. Por desgracia, su calidad no estaba a la altura de su cantidad. Las monedas de los últimos siglos de Cartago, particularmente las del período vándalo, muy pequeñas y pobremente acuñadas, están también gravemente corroídas por su permanen-

RESIDENCIA ROMANA reconstruida más tarde respetando la misma planta. La está excavando un equipo de la Universidad de Michigan en Cartago. Situada en una de las muchas terrazas que los romanos construyeron para cubrir las ruinas de la Cartago púnica, la primera casa fue edificada en este lugar a mediados del siglo primero d.C. Poco antes del 400 d.C. se reconstruyó como la vemos aquí. El mosaico que representa a los aurigas griegos estaba en el umbral que conducía del comedor al patio. Las grandes zanjas *(color)* fueron abiertas más tarde por ladrones de materiales.



CAMPAÑA DE EXCAVACIONES de urgencia. Estas son, de norte a sur, un yacimiento al norte de las Termas Antoninianas (danesa), una iglesia cristiana (búlgara), una exploración de las calles y de la muralla (italiana), villas romanas (tunecina), la ciudadela púnica (francesa), un yacimiento al sur de las termas (alemana), la zona de la Universidad de Michigan, una exploración cerca del circo (polaca), dos yacimientos cerca del puerto militar y uno en la muralla de la ciudad (británica) y una concesión conjunta de la Universidad de Harvard y de la Universidad de Chicago.

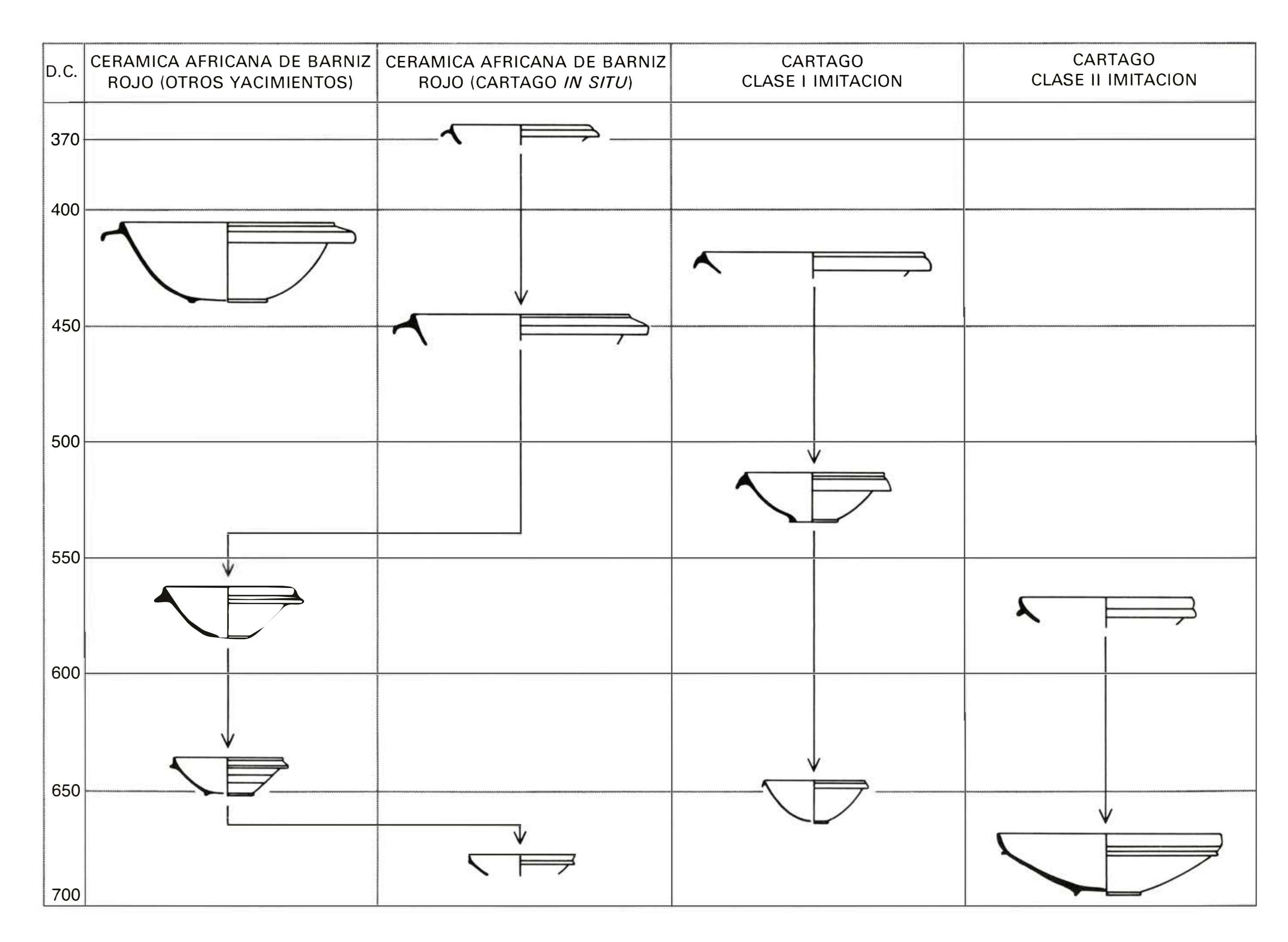
cia en un suelo ácido. Otra ventaja más del tamizado, como más adelante se verá, fue la recogida de una gran cantidad de huesos de animales y de materiales de plantas.

Una vez que se ha establecido un encuadre cronológico seguro, se hace posible interpretar las sucesivas fases de ocupación del yacimiento en relación con los anales de la historia clásica que han sobrevivido. Por ejemplo, ahora podemos ofrecer, por lo menos a grandes rasgos, la historia de la zona que hemos excavado hasta el presente. Nuestro trabajo comenzó en 1975 limpiando una zona de unos 15 por 25 metros en el lado noroeste de nuestro campo de 100 por 400 metros. La excavación puso al descubierto gran parte de la planta baja de una casa romana que daba a una de las calles normales orientadas de norte a sur en Colonia Julia. Los reconstructores romanos de Cartago habían levantado una terraza en esta parte de la ciudad acumulando una gruesa capa de relleno sobre los escombros subyacentes de la Cartago púnica. Todavía nos faltan por excavar los niveles púnicos. Aquí, como es evidente en otras partes de la ciudad, la erección de terrazas sobre las ruinas púnicas se verificó antes de trazar la cuadrícula de las calles principales y las subdivisiones secundarias de la misma, es decir, las *insulae* o manzanas. La calle principal, a que nuestra casa da frente, se denomina Kardo IX Este en el callejero de la Cartago romana.

La fecha de la construcción de la terraza cae a principios del siglo primero d.C. Parece, por tanto, que esta zona, situada al oeste del puerto circular, no figuró entre las primeras partes de la ciudad que fueron reconstruidas cuando Julio César y su sucesor Augusto establecieron la Colonia Julia. En todo caso quienes planificaron la ciudad subdividieron las nuevas manzanas en 16 parcelas iguales, cada una de ellas de unos 17,5 metros cuadrados. La más antigua de las casas romanas de nuestro sector, edificada no después de mediado el siglo primero, ocupaba una de esas parcelas o como mucho dos.

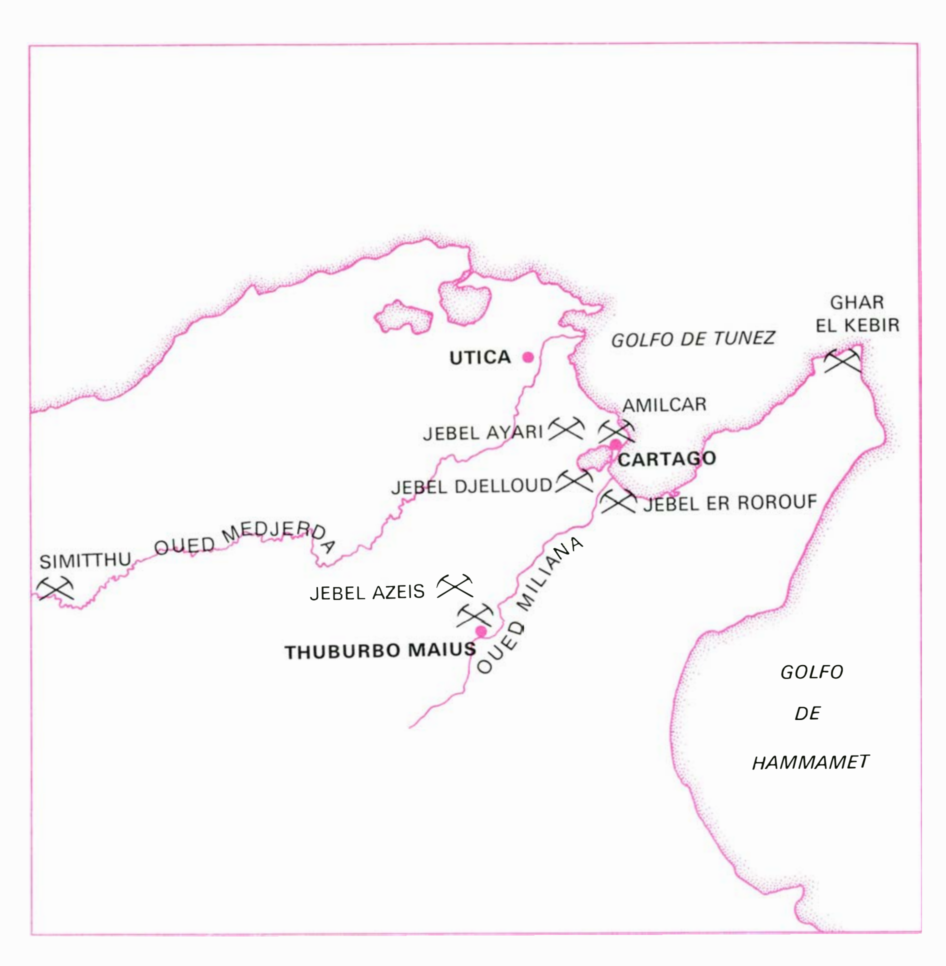
No hemos excavado todos los niveles romanos primitivos en el solar de la casa. Lo que hemos puesto al descubierto demuestra que una cisterna grande fue construida bajo el patio de la casa a comienzos del siglo segundo. Hacia mediados de ese mismo siglo fueron abiertos unos hoyos al borde de la calle, inmediatamente por fuera de la casa; poco después fueron rellenados y cubiertos por una acera pavimentada. Si esta obra realizada en la calle señala una ruptura en la actividad de la ocupación, pudiera estar relacionada con un gran incendio que asoló Cartago durante el reinado del emperador Antonino Pío (138-61 d.C.) y destruyó una gran parte de este sector de la ciudad.

Nuestra tarea se ha concentrado en los restos de la casa que se pueden atribuir a los períodos romano tardío, ván-



EVOLUCION DEL ESTILO de la cerámica durante unos 300 años en una especie de cuenco con reborde que comenzó a adquirir popularidad a fines de la época romana. Era una cerámica fina de barniz rojo, copiada a través del Imperio. Ejemplares hechos en Africa *(1.ª y 2.ª columnas)* **se encuentran en Cartago y en otros sitios. Los cuencos fueron imitados por ceramistas de Cartago y aledaños** *(3.ª y 4.ª columnas).*

dalo y bizantino de Cartago. En relación con el período romano tardío, es evidente que la casa fue completamente reconstruida poco antes del 400 d.C. También pudo haber sido ampliada al mismo tiempo, aunque parece que sólo fue ocupada la mitad del ancho de la manzana. La principal habitación de la casa, de siete por nueve metros aproximadamente, era un comedor; la pared oriental de este comedor, correspondiente al lado de la calle, está situada a cinco metros al oeste del muro de la fachada. Por el norte, el comedor se abría a un patio columnado, y la habitación misma estaba pavimentada con un mosaico figurado. Los ocupantes de la casa tras el intermedio vándalo de Cartago reemplazaron el mosaico romano original, pero parte del borde de este mosaico, con un dibujo de roleos de acanto, logró sobrevivir.



LAS CANTERAS DE CARTAGO comprendían una en las afueras de la ciudad y tres lo bastante cerca de la costa para permitir el transporte por mar. Otras canteras importantes estaban situadas entre distancias de 12 (el yeso de Jebel Ayari) y 160 kilómetros (el mármol numídico de Simitthu).

Sucede que en los siglos siguientes al 100 a.C., la Cartago romana fue la sede de una de las más importantes escuelas de confección de mosaicos dentro del Imperio. La exploración tunecina de nuestro campo descubrió un ejemplar muy bien conservado. En 1975 pudimos comprobar que ese mosaico había ocupado el umbral situado entre el comedor y el patio de nuestra casa. En él están representados cuatro aurigas griegos en sus respectivos carros a las puertas de salida de un circo *(véase ilustración de la página 92)*. Un pequeño estanque ocupaba el ángulo suroeste de la columnata del patio; tanto el patio como el pórtico estaban pavimentados de mosaicos, pero sólo algunos fragmentos de estas zonas pavimentadas se han librado de la destrucción.

Donde estos y otros fragmentos de mosaicos sobreviven nos dicen algo sobre los ocupantes de la casa durante la dominación de los vándalos. Katherin Dunbabin, de la Universidad de McMaster, ha estudiado todos nuestros mosaicos. Gracias a ella sabemos que conforme los mosaicos romanos se estropeaban o pasaban de moda, los ocupantes de la casa los reemplazaban por otros nuevos e incluso ampliaban las zonas pavimentadas. Por ejemplo, la reconstrucción efectuada durante el período vándalo amplió la casa hacia el este, hasta incluir la zona de la calle que había sido anteriormente el firme de la acera. La obra comportó la ampliación de las habitaciones a lo largo del lado oriental de la casa, y al mismo tiempo los pisos de las habitaciones ampliadas fueron recubiertos de nuevos mosaicos. Exceptuadas estas reparaciones y modificaciones, el siglo aproximado de dominio vándalo en Cartago parece haber afectado muy poco a la que ahora llamamos la Casa de los Aurigas Griegos.

Poco después de que el general bizantino Belisario expulsase de Cartago a los vándalos, nuestra casa experimentó una reconstrucción a fondo. La obra se atuvo al plano y a las paredes existentes, pero el piso del comedor volvió a ser pavimentado con placas de mármol de un diseño muy complejo. Los pavimentos de esta clase se conocen como *opus sectile* (la expresión latina para obra de recortes) y el ejemplo de nuestra casa es el primer *opus sectile* bizantino encontrado en una casa particular de Cartago y que se puede fechar con confianza. Las placas de mármol han sido estudiadas por nuestro geólogo, Reuben G. Bullard. Según su informe, muchas de ellas no fueron cortadas expresamente para el fin que cumplieron, sino que fueron recogidas de pavimentos similares en otros sitios y reutilizadas aquí.

La obra bizantina en la casa supuso la conversión de la zona comprendida entre el comedor y la calle en un patio abierto y pavimentado con losas. El resto de las habitaciones que daban a la calle recibieron nuevos mosaicos, pero desde entonces la casa no fue objeto de otros cambios hasta el tercer cuarto del siglo sexto. Fue entonces cuando el patio que daba frente al comedor, y que tenía un piso de cemento en aquella época, fue cubierto también de losas. La casa parece haber estado ocupada hasta la conquista árabe del 698 d.C., pero sin que le pusieran nuevos mosaicos. Sus habitantes intentaron conservar el pavimento de *opus sectile* del comedor. Dos arreglos de distintas épocas se aprecian en él. El primero se ajustó en todo lo posible al dibujo original; el segundo renunció a adaptarse y los arreglos estropearon el diseño.

Durante la construcción del ferrocarril de Túnez en 1909, en las proximidades de nuestro yacimiento, los obreros que construyeron el arcén con derecho de paso arrasaron los niveles superiores de nuestra casa para utilizarlos como material de relleno. En consecuencia, es muy poco lo que queda de los restos de

habitación bizantina tardía y postbizantina. Afortunadamente, existen vestigios de este período de la historia cartaginesa en otras zonas de nuestra parcela, justamente al sur de la iglesia y del baptisterio en ruinas que habían sido localizados anteriormente por el Instituto Nacional tunecino. Toda la campaña de 1976 la hemos dedicado a trabajar en esta parte de la parcela, y descubrimos en ella testimonios de cinco fases sucesivas de ocupación, que abarcan desde el 650 d.C., aproximadamente, hasta el siglo XI poco más o menos.

Durante la primera de las cinco fases, la zona había formado parte de un activo conjunto eclesiástico que comprendía la iglesia inmediata a nuestra casa. Nuestra cronología indica que el templo y las construcciones dependientes del mismo alcanzaron su fase final algo antes de 647-59 d.C. La fase siguiente, fechada algo después de 659-68 d.C., señala el aparente derrumbamiento del conjunto de la iglesia. Este colapso viene indicado por un proceso con el que los arqueólogos clásicos están muy familiarizados: el robo de materiales, esto es, el despojo de piedras de construcción en un yacimiento abandonado para su empleo en otro lugar.

La tercera fase, datada algo después de 668-73 d.C., ofrece señales de tres esfuerzos diferentes, por lo menos, para renovar la pavimentación de la ruina abandonada, todos ellos de pobre calidad. Es significativo que la segunda y la tercera fase del yacimiento, caracterizadas en conjunto por unas condiciones en Cartago que nada tienen de prósperas ni de estables, se adelanten muchos años a la destrucción musulmana de la ciudad en 698 d.C. Diríase que el conflicto permanente entre el Islam y el Imperio Bizantino, que comienza medio siglo antes en el Próximo Oriente, también había desestabilizado por estas fechas las provincias norteafricanas.

La cuarta fase de nuestra precedente iglesia, fechada con posterioridad a 687-95 d.C. y quizá llegando hasta el siglo X, se distingue por una segunda y más intensa incursión de ladrones de materiales. La fase quinta y final puede situarse con ciertas garantías en el período islámico del siglo X o incluso del XI. Uno de los pisos de esta fase se sobrepuso y selló una moneda islámica, lo que permite conjeturar que el piso es posterior en fecha a la moneda.

Normalmente la arqueología clásica presta poca atención a un yacimiento en su fase postrera, una vez que los edificios monumentales han caído en desuso. Es probable que las ocupaciones posteriores se caractericen por una construcción menos sólida que la cantería, como pueden serlo los tabiques de adobe o de cañizo. Los trabajos arqueológicos últimos en Inglaterra han revelado parte de lo que ocurrió allí al término de la ocupación romana de Britannia, pero son raros los trabajos de este género en otras partes del Imperio Romano. Por consiguiente, nuestros hallazgos en este rincón de la Cartago postpúnica arrojan una valiosa luz sobre los últimos días de la ciudad.

En pocas palabras, aun después de que los tejados se habían derrumbado y los muros habían sido despojados de sus piedras, el cascarón del conjunto de la iglesia albergó una serie de habitaciones domésticas en las que todavía podemos reconocer unidades-casa individuales. Las condiciones eran precarias, los tabiques endebles y los suelos no estaban pavimentados de mosaicos ni de una mezcla fina, sino de adobe apisonado. Aun así, la cantidad de monedas y tiestos aparecidos en los niveles superiores indican la intensa actividad que caracterizó los últimos días de Colonia Justiniana. El avance del Islam, en efecto, había dejado a Cartago cercada. Es evidente que los cristianos de la campiña circundante que se refugiaban en la ciudad agradecían cualquier cobijo por rudimentario que fuese.

Los restos de animales (que hemos recuperado por tamizado en seco) y los de plantas (recogidos de la tierra tamizada, por flotación) dan una dimensión más a nuestro cuadro de la vida en la Cartago postpúnica. Los huesos y las conchas no han sido contemplados normalmente con mucho entusiasmo por los arqueólogos clásicos. Nosotros hemos tenido la fortuna de contar con David Reese, estudiante graduado de la Universidad de Cambrigde, para el análisis de nuestros hallazgos de estos materiales. El descubrió que los cerdos y las ovejas o cabras eran los animales de carne más abundantes en la ciudad. Los huesos de vacuno son menos corrientes, lo que significa que estos animales mayores eran uncidos al arado o al carro antes que sacrificados para carne. Esta preferencia se mantiene inalterable en Cartago desde los comienzos del período romano hasta el bizantino tardío, y el hecho contrasta fuertemente con las fuentes literarias latinas que pretenden hacernos creer que la carne de vaca era la más popular. La presencia del camello con anterioridad al influjo árabe está comprobada por el hallazgo de huesos de este animal en un hoyo lleno de basura poco posterior al 100 d.C.

Espinas de pescados representantes de varias especies indican que la fauna marina constituía una parte importante de la alimentación cartaginesa. Es curioso que el atún, el pescado de mayor consumo en la actual cocina de Túnez, no aparezca por ningún lado. Hay gran variedad de conchas de moluscos, tanto marinos como terrestres. Sin embargo, no se encuentran pruebas de que se comiesen los caracoles terrestres; las señales de desgaste por agua indican que los caracoles y los bivalvos marinos tampoco eran consumidos, aunque fuesen recogidos accidentalmente junto con la arena de playa que se empleaba para hacer cemento.

Richard I. Ford y Naomi Miller, de la Universidad de Michigan, han analizado los materiales botánicos del yacimiento, que incluyen una serie de semillas y de madera quemada. Los huesos de aceituna son las semillas más abundantes. Quienes comían estas aceitunas no eran sólo los varios habitantes de nuestra parcela, sino también los obreros que preparaban el mortero para los canteros; los huesos de esas aceitunas aparecen a menudo en el cemento de las construcciones de todas las épocas. Otras semillas identificadas por Ford y Miller consisten en granos de trigo, pepitas de uvas y el hueso de un melocotón. Al mismo tiempo, se ha formado una colección de madera de los árboles que crecen hoy en Túnez y se ha buscado información en los autores clásicos acerca de las principales fuentes de madera en la antigüedad, y de las clases de la misma que eran preferidas para distintos fines. El siguiente paso consistirá en ver cómo la madera quemada que hemos recogido en nuestras excavaciones encaja dentro de ese panorama teórico.

Un esfuerzo análogo se ha realizado respecto a las diversas piedras de construcción usadas en Cartago. Lo mismo los alrededores inmediatos de la ciudad que zonas más distantes de Túnez, e incluso de Argelia, son bien conocidas por las canteras abandonadas de las épocas púnica y romana. Lo que ya no era tan conocido, sin embargo, es cómo se empleaban los distintos productos de las canteras.

Bullard, nuestro geólogo, ha estudiado el problema. Según él los romanos siguieron las costumbres púnicas en relación con una arenisca basta y porosa extraída de Ghar el Kebir, una cantera situada a unos 60 kilómetros al nordes-

te de Cartago. Esta piedra, que se podía traer por mar a la ciudad, era usada como material general de construcción. Los romanos empleaban como material de relleno, más que como cantería exterior, una arena local petrificada y una arenisca cuarzosa de las canteras de Amílcar situadas al norte de la ciudad. Una caliza duradera y fácil de trabajar podía obtenerse en las canteras de Jebel Djelloud y de Jebel er Rorouf, de donde la piedra había de ser transportada también por mar, y de Jebel Ayari, situada a unos 12 kilómetros tierra adentro. Los romanos empleaban esta excelente piedra para pilares, dinteles y losas de pavimento, y, en alguna ocasión, también para hiladas de los muros.

Para revestimientos de mayores pretensiones los romanos explotaban una piedra próxima al mármol, gris con venas rosadas y blancas, de un afloramiento de sedimentos jurásicos situado a unos 60 kilómetros al suroeste de Cartago, cerca de Thuburbo Maius. Para columnas grandes seleccionaban una caliza de venas blancas y pardas de un depósito adyacente. Las arcillas ricas en yeso de un depósito del Jebel Azeis, al norte de Thuburbo Maius, y la cantidad de yeso accesible aun más cerca de Cartago explican el frecuente uso de bloques de yeso moldeado en la construcción urbana. La cantera de mármol de Simitthu, a 160 kilómetros al oeste de Cartago, fue para los romanos la fuente del *marmor Numidicum,* una piedra de revestimiento rosa y amarilla que no sólo se empleó en la ciudad sino que también se exportó en grandes cantidades a otras ciudades en toda la extensión del Imperio.

Conforme adelanta nuestro trabajo esperamos enlazar otras canteras tunecinas con las piedras de la construcción de la Cartago postpúnica, de la misma forma que nuestros otros colaboradores de diversas especialidades van aportando las contribuciones de sus campos.

El ejemplo más inmediato es el trabajo realizado el año pasado en nuestra parcela de 100 por 400 metros. Exploraciones realizadas mediante un contador de resistencia eléctrica en la mitad meridional del campo han suministrado importante información sobre la presencia y ausencia de cisternas, información que permitirá juicios de valor con respecto a temas más amplios como pueden ser el desarrollo arquitectónico, la planificación urbana y la densidad de población. Junto con nuestros colegas en este esfuerzo internacional de urgencia vamos ampliando cada campaña los límites tradicionales de la Arqueología clásica.

Juegos matemáticos

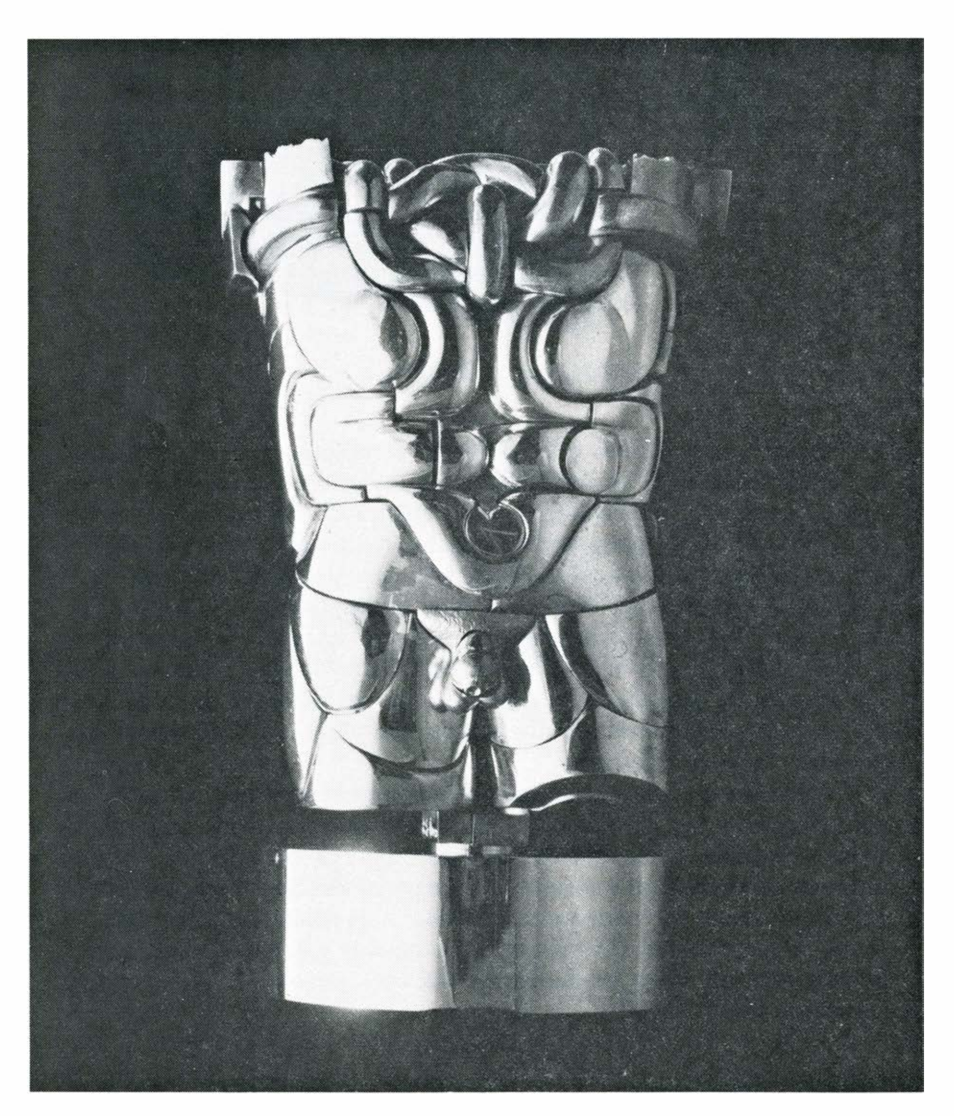
Todas las esculturas de Berrocal son desmontables, lo mismo que los rompecabezas mecánicos de enclavamiento

Martin Gardner

"Nos embarcamos después en una disquisición sobre la existencia de formas de belleza desprovistas de utilidad. El General mantuvo que eran imposibles. El Dr. Johnson mantuvo que sí lo eran, y puso como ejemplo la taza de café que sostenía en la mano, cuya decoración carecía de utilidad real, pues la taza contendría el café exactamente lo mismo si no estuviese decorada; no obstante, el diseño era hermoso".

JAMES BOSWELL,
Life of Samuel Johnson

El artículo de este mes se propone presentar a Miguel Berrocal a los lectores, quizás el más sobresaliente de los escultores españoles vivos.



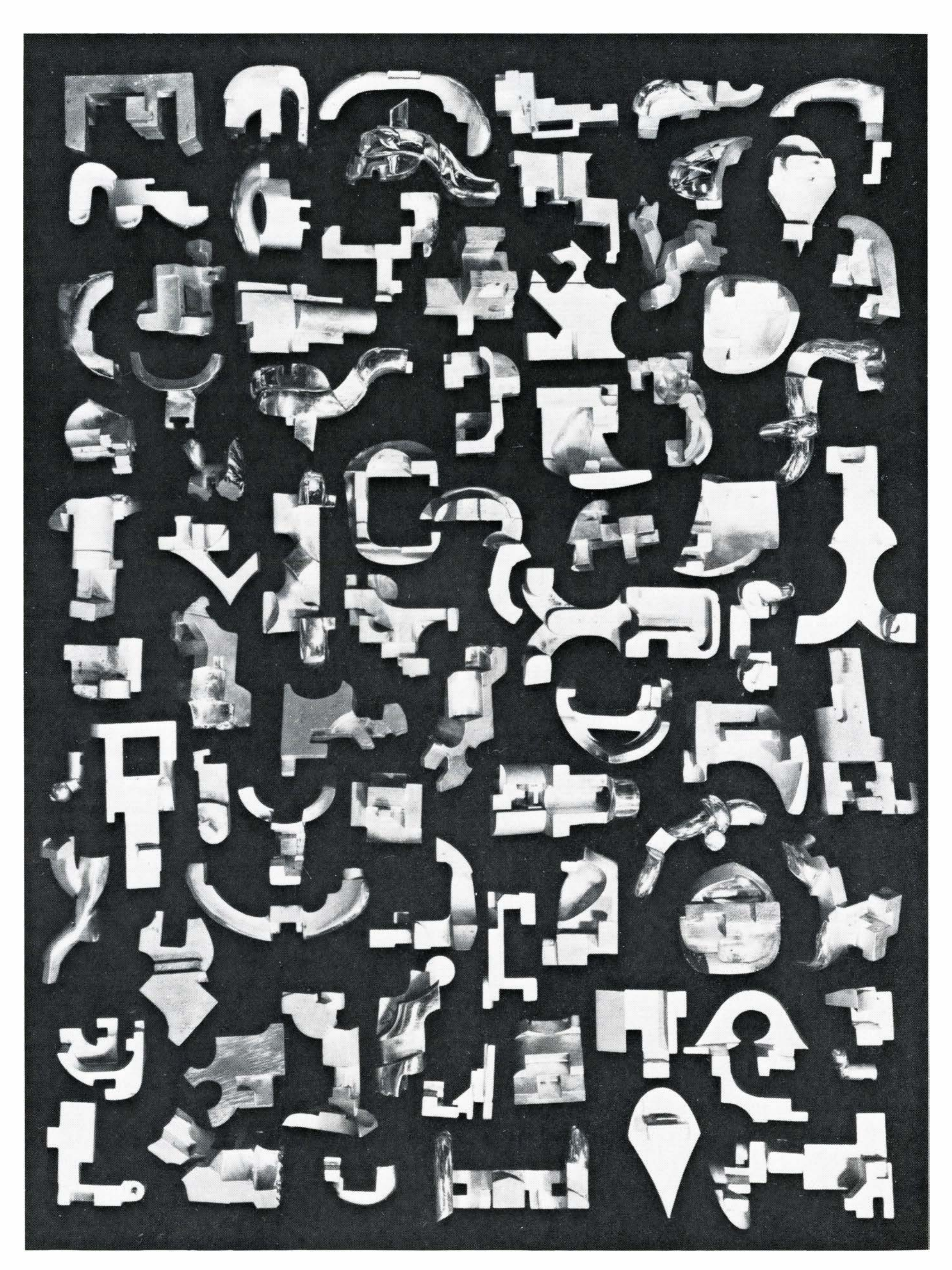
Goliath, *de Miguel Berrocal*

En Europa, Berrocal es ya objeto de culto, y sus adeptos crecen constantemente. No obstante, antes de comentar con más detalle los notables hallazgos de Berrocal, y de explicar por qué debe analizarse su escultura en una sección de matemática recreativa, es preciso echar una ojeada a una de las categorías más antiguas de rompecabezas mecánicos.

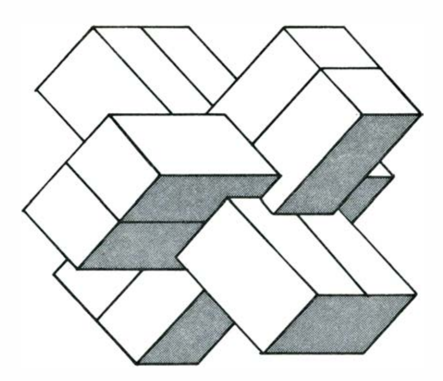
¿Es posible que alguno de los lectores no haya tenido nunca en sus manos uno de esos montajes de madera, formados por piezas tan hábilmente ajustadas que resulta casi imposible desmontarlos? Tales objetos suelen llamarse rompecabezas chinos, y, por lo general, una vez desmontados, todavía resulta más difícil volver a ensamblarlos. Suelen poseer una pieza especial, llamada pieza maestra o pieza clave, que debe retirarse en primer lugar para poder desmontar las restantes. Para recomponer muchos de estos rompecabezas es necesario ir colocando las piezas en una cierta secuencia, insertándose la clave en último lugar, y así enclavar firmemente en su sitio a las demás partes. Se han vendido en todo el mundo centenares de juegos de este tipo durante siglos, muchos de ellos de inventor anónimo (aunque también se han expedido cientos de patentes con variantes no comercializadas de tales rompecabezas).

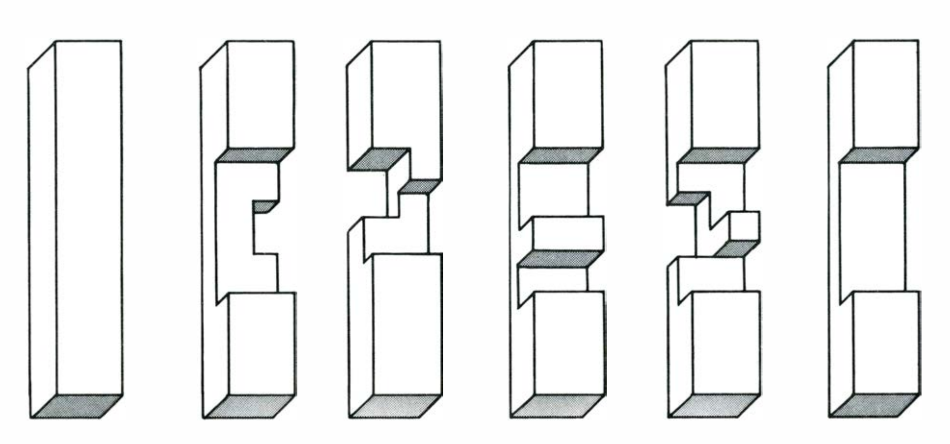
Apenas se sabe nada de los albores históricos de estos montajes, pero seguramente el de "rompecabezas chino" sea nombre poco apropiado. Sin duda, los países de Asia oriental producían rompecabezas de esta clase ya en el siglo XVIII, pero lo mismo ocurría en los países de Europa. Además, rompecabezas mecánicos de otros tipos, de incuestionable origen europeo, han sido denominados también "rompecabezas chinos". Quizá, como ha escrito Joseph Needham en el volumen III de su *Science and Civilization in China:* "Tal vez los europeos se sintieran inclinados a dar a los enigmas mecánicos el nombre de la que, para ellos, era una civilización enigmática".

Hay rompecabezas chinos compuestos por tres piezas solamente, pero el más sencillo de los modelos no triviales es el muy conocido de seis piezas que se muestra en la ilustración superior de la página 104; debajo del modelo montado se muestran las seis piezas que lo forman, tal como fueron dibujadas por el "Profesor Hoffmann" (seudónimo de Angelo Lewis) en su libro sobre enigmas mecánicos, *Puzzles Old and New,* publicado en 1893. La barra íntegra, en el extremo



Las 80 piezas de **Goliath**





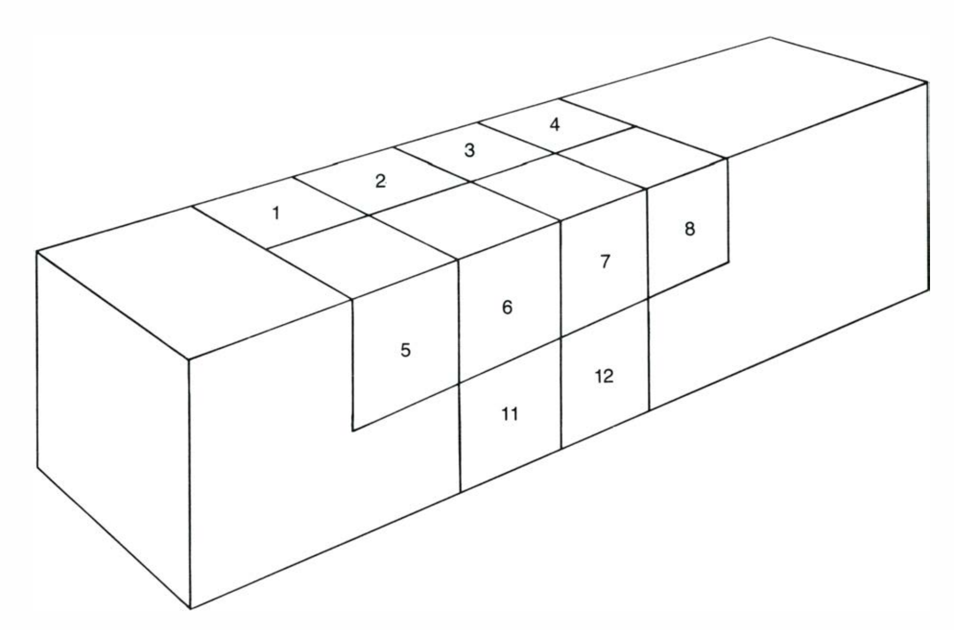
Rompecabezas "cardo" de seis piezas

izquierdo de la ilustración, es la clave. Otro sistema de seis piezas, distinto del anterior, apareció ya en *The Magician's Own Book,* obra de autor anónimo, que se publicó en los Estados Unidos por primera vez en 1857. Anthony S. Filipiak, en su obra *100 Puzzles: How to Make and Solve Them* (A. S. Barnes and Co., 1942), llama "cardos de seis piezas" a tales montajes, presumiblemente, porque su forma los recuerda. En la actualidad, y al menos en los Estados Unidos, todos los rompecabezas de esta categoría suelen denominarse "cardos", por extensión.

Las distintas versiones de "cardos" manufacturadas pocas veces son parecidas, lo que sugirió a Filipiak un problema sorprendentemente difícil de geometría combinatoria. Imaginemos que la sección central de la pieza maestra (la pieza clave no ranurada), se divide en 12 cubos unidad, como se muestra en la ilustración inferior de esta página. El lado de cada uno de estos cubos es la mitad del grosor de la pieza maestra. Hay cuatro cubos más, situados bajo los numerados 2, 3, 6 y 7, pero solamente dos de ellos, el 11 y el 12, son visibles en la ilustración. Cada una de las piezas de cualquier cardo de seis queda completamente descrita especificando cuáles de los 12 cubos unidad se han suprimido para obtenerla.

Como cada uno de los cubos unidad, o bien forma parte de una pieza del rompecabezas, o bien ha sido suprimido para construirla, en principio habría 2^{12}, esto es, 4096, posibles formas para las piezas del cardo. Evidentemente, es preciso eliminar aquellas que dividan la barra en dos trozos. Entre las restantes, se comprueba que cada forma se repite ocho veces, por cuatro rotaciones simples de eje longitudinal y dos medias vueltas de eje transversal, que permutan entre sí las posiciones de los extremos. (No se suprimen, en cambio, las imágenes especulares de configuraciones asimétricas). Finalmente, hay ciertos modelos que por razones mecánicas no pueden formar figuras con enclavamiento al combinarlas con cualesquiera otras. Estas configuraciones pueden desecharse también. Una vez eliminadas todas las configuraciones aludidas, ¿cuántas formas pueden tener las seis piezas componentes del cardo? En opinión de Filipiak, su número era 432.

Más recientemente, han sido preparados programas de ordenador para analizar mecánicamente el problema por tres matemáticos, William H. Cutler, del Wartburg College de Waverly, Iowa, Robert H. Mackay, de Londres, y C. Arthur Cross, de Cheshire, Inglaterra. Cutler y Cross coinciden en afirmar que el número de piezas utilizables es 369. No obstante, sesenta y siete de ellas solamente pueden utilizarse con duplicados suyos, y dos solamente con triplicados, lo que hace un total de 440 piezas. Se plantea así una segunda cuestión, todavía no resuelta: ¿Cuántos distintos cardos de seis piezas pueden construirse con el anterior conjunto? Es necesario desechar las combinaciones de piezas que, o no llegan a enclavarse, o solamente forman sistemas totalmente bloqueados, que es imposible desmontar. También es recomendable suprimir las construcciones que dejen huecos internos, que



Configuraciones de las piezas de un "cardo"

por lo común no se sostienen bien. Con estas restricciones, el programa de Cutler halló que el número de distintos cardos de seis piezas es de 119.963, aunque Cutler no está completamente satisfecho de su programa, que teme pueda contener errores. Los lectores interesados en la cuestión pueden escribirle solicitando detalles.

Solamente conozco dos sitios donde puedan obtenerse actualmente tipos más complicados de rompecabezas "cardo". En Inglaterra, Paradox Engineering Limited (Over Wallop, Stockbridge, Hants, England) envía bajo demanda catálogos de su serie "Pentangle", en la que están incluidos varios estupendos ejemplares de cardos realizados en madera. El llamado "Abuelo Chuck" (Grandpa Chuck) consta de 96 piezas. En los Estados Unidos, enviándole a Stewart T. Coffin (Old Sudbury Road, R.F.D. 1, Lincoln, Mass., 01773) un sobre comercial ya franqueado y dirigido a uno mismo, se recibirá un folleto de su insólita serie de rompecabezas originales de tipo cardo. Sus precios son elevados, ya que se trata de piezas de artesanía, ajustadas a mano, realizadas en madera dura. Durante varios años Coffin ha estado publicando intermitentemente un boletín, *Puzzle Craft*, que trataba de la historia y la construcción de rompecabezas mecánicos. Según me dice, enviará una reimpresión de los seis números publicados, 24 páginas en total, juntamente con el catálogo, contra reembolso de 2 dólares.

De vez en cuando se han ido fabricando rompecabezas chinos que recuerdan objetos familiares: un coche, una pistola, un buque de guerra, un avión, una pagoda, un tonel, un huevo y diversos animales. La mayoría de los rompecabezas autoblocantes no se diseñan con la intención de que sus formas recuerden objetos concretos; no obstante, por sus simetrías siempre es agradable contemplarlos. Por lo general, los modelos que podríamos llamar "figurativos" son, estéticamente al menos, poco atractivos. Lo que nos lleva al trabajo de Berrocal. Que yo sepa, ha sido el primero que ha sabido combinar los rompecabezas chinos de enclavamiento y un trabajo artístico de la mayor calidad.

Berrocal nació en 1933, en una familia malagueña de clase media. Estudió arquitectura y matemáticas en la Universidad de Madrid, y, posteriormente, arte en París y Roma, acabando por establecerse en Negrar, en las afueras de Verona. En la actualidad reside allí, en una villa palaciega, con su segunda esposa, la princesa Cristina, nieta del último rey de Portugal. Berrocal preside una

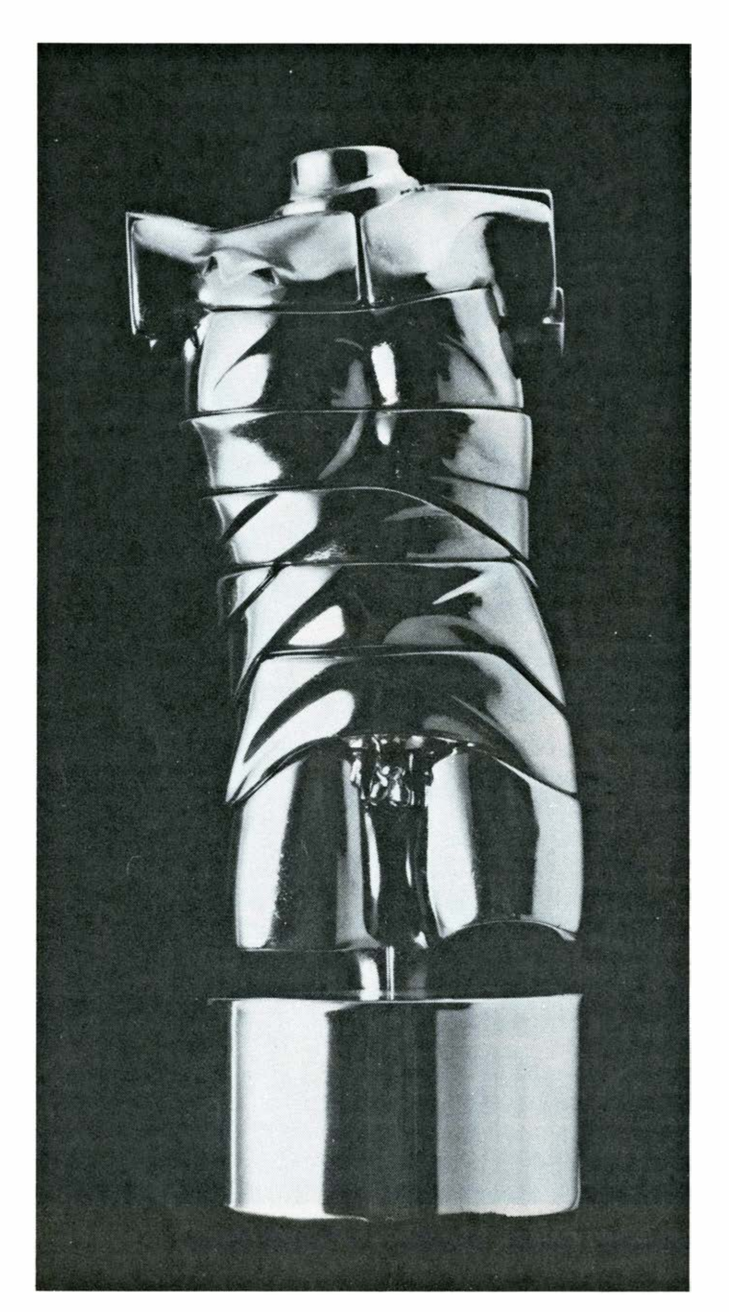


Portrait de Michèle (cabeza) *acoplado a* **La Totoche (cuerpo)**

fundición en Negrar, con más de 200 operarios, en la que no solamente se vacían sus propias obras, sino también muchos trabajos de otros escultores europeos. "Soy el patrón de la Mafia de escultores", dijo en cierta ocasión. Lo mismo que Pablo Picasso y Salvador Dalí, los más notables pintores españoles del siglo XX, Berrocal es un virtuoso que ha sabido combinar una prodigiosa producción con una hábil política de relaciones públicas y está dotado de una personalidad inmodesta y deslumbrante.

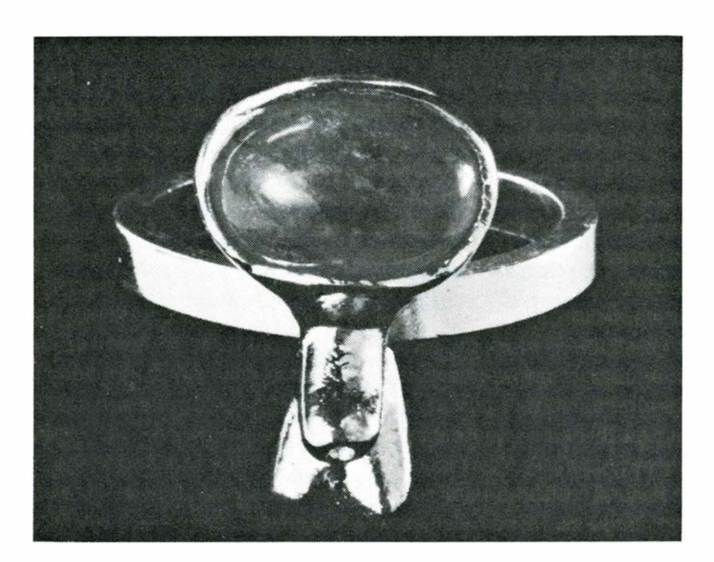
Es imposible apreciar la combinación única de valores contenidos en la obra de Berrocal –belleza visual, placer táctil, humor, el reto intelectual de un rompecabezas combinatorio tridimensional– si no se ha conseguido antes desmontar y volver a montar varias veces alguna de sus obras. Consideremos, por ejemplo, los dos Berrocales que se muestran, superpuestos, en la ilustración superior. La cabeza, llamada *Portrait de Michèle* (Retrato de Michèle) está compuesta por 17 elementos distintos, de formas curiosas, cada uno de ellos diseñado para ser individualmente una escultura abstracta, y de forma que resulte agradable de manipular. El cuerpo, llamado *La Totoche* (La rellenita) es uno de las varios cuerpos a los que puede acoplarse la cabeza; puede desomponerse en 12 piezas.

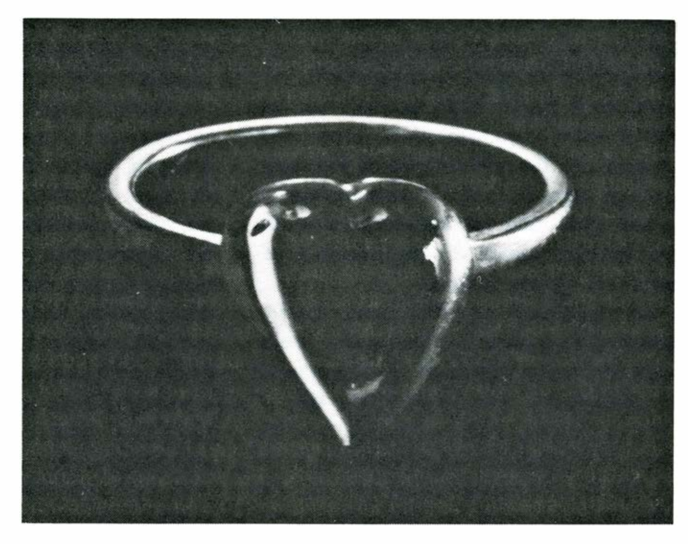
Berrocal creó el nombre de "múltiplos" para las copias, que fabrica en grandes cantidades, de sus obras. Su edición del *Portrait de Michèle* es típica: seis múltiplos en oro macizo, 500 en plata fina, y 9500 en bronce niquelado. Cada pieza está firmada y numerada, impecablemente terminada con la precisión de un trabajo de ingeniería. En todos los Berrocales, las piezas se desmontan y se ensamblan en una cierta secuencia. Para





Mini David (izquierda) *y* **Mini Cariatide** (derecha)





Anillos pertenecientes a **Mini David** (**izquierda**) *y a* **Mini Cariatide** (**derecha**)

desmontar el *Portrait de Michèle* es necesario extraer primero un elemento del cuello. Los rompecabezas chinos tienden a desmantelarse solos en cuanto se les retiran algunas piezas, pero un Berrocal se mantiene firmemente unido hasta que se han separado sus dos últimas piezas. En muchos casos, el modelo montado es macizo, es decir, no tiene huecos interiores, y hasta haber separado la pieza n es imposible retirar la $n+1$. Los enclavamientos son tan ingeniosos, que a veces es imposible sacar una pieza sin alterar ligeramente las posiciones de las restantes. Cada "múltiplo" viene acompañado de un libro de instrucciones encuadernado, ilustrado por Berrocal. Cada etapa del montaje está dibujada en una hoja distinta, en la que se muestra en color la pieza a separar. Al final, un dibujo en perspectiva isométrica exhibe el perfil de todas las piezas (en transparencia), ya situadas en su sitio. No obstante, aún consultando el manual de instrucciones, pueden hacer falta varios días para dominar la técnica de desmontar y ensamblar un Berrocal.

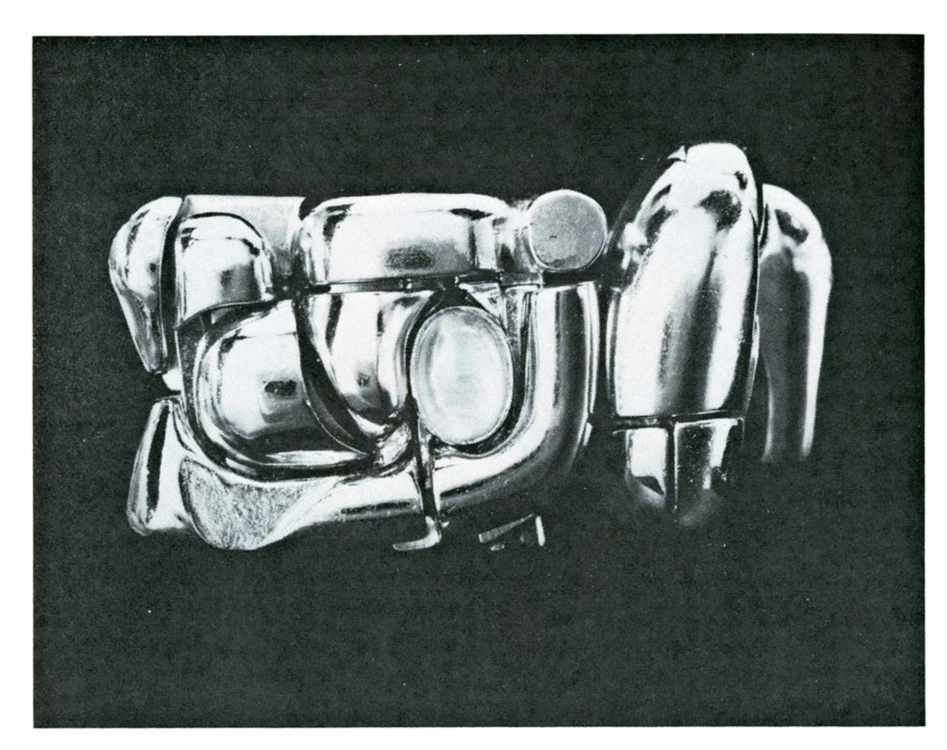
El número de piezas de los Berrocales varía desde solamente tres hasta cerca de 100. Exquisitos anillos y brazaletes, todos ellos utilizables para el atuendo personal, forman parte de muchas de sus esculturas. Por ejemplo, la pupila del ojo de Michèle es la piedra aguamarina de un anillo. Otro popular Berrocal es *Mini David,* torso mostrado en la ilustración superior de la página opuesta. Una de sus 22 piezas es el anillo que se muestra en la ilustración inferior de la página opuesta. Los genitales del torso cuelgan de la gema del anillo. La edición completa de esta obra ha sido vendida. Está formada por seis múltiplos de oro macizo, 500 de bronce chapado en oro y 9500 de bronce niquelado. En los de oro, la gema del anillo es de jade verde; en los de bronce niquelado es zafiro. Frecuentemente, Berrocal diseña versiones "micro" de sus obras para ser utilizadas como pendientes. Hay un *Micro David* que contiene un anillo con banda de malla y un corazón azul de lapislázuli. La estructura interna de una obra micro es siempre totalmente diferente de su contrapartida "mini".

Mini María, que se muestra en la ilustración superior de esta página, solamente puede desmontarse presionando sobre un cojinete de bolas de una pierna. La escultura está compuesta por 23 piezas. Una de ellas es un anillo con una adularia, que forma uno de los pechos de María. Dentro de la figura hay un órgano sexual masculino, que se ve en la ilustración inferior de esta página, y que se descompone en cinco piezas, dos de las cuales son bolas de acero. El correspondiente elemento del colgante *Micro María* tiene una diminuta aguamarina en el extremo del órgano. La gema está adosada a un anillo con una banda de malla.

Otra figura reclinada, *Mini Zoriade* se abre al girar un zapato. Los pechos de Zoriade son las adularias de un anillo. *Mini Cariatide* se muestra en la derecha de la ilustración superior de la página opuesta. El monte de Venus de la figura se encuentra en el anillo que se muestra en la parte derecha de la ilustración inferior de la página opuesta.

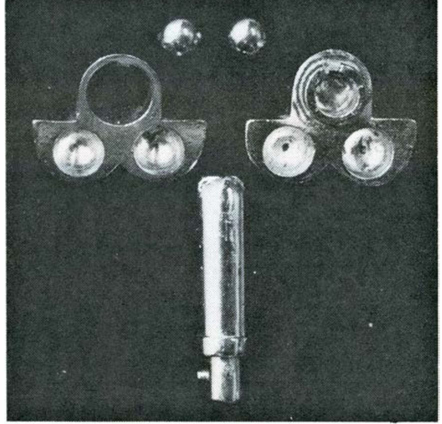
Goliath, que se muestra en la ilustración de la página 102, es el más complicado de los trabajos de Berrocal. La totalidad de sus 80 elementos aparecen en la ilustración de la página 103. La escultura se muestra parcialmente desmontada en la ilustración superior de la página 108, para poner de manifiesto cómo un Berrocal es una extraordinaria obra de escultura abstracta en cualquier estadio de ensamblaje. Cuando el torso de Goliath está completamente montado, se puede girar su hoja de parra para dejar al descubierto los genitales. Existen en realidad dos piezas que representan los genitales, una circuncidada y la otra no, y el torso puede ensamblarse con cualquiera de ellas al descubierto.

El féretro que Berrocal preparó para Romeo y Julieta se ve en la ilustración inferior de la página 108, juntamente con una obra anterior, de 16 piezas, que retrata entrelazados a los malhadados

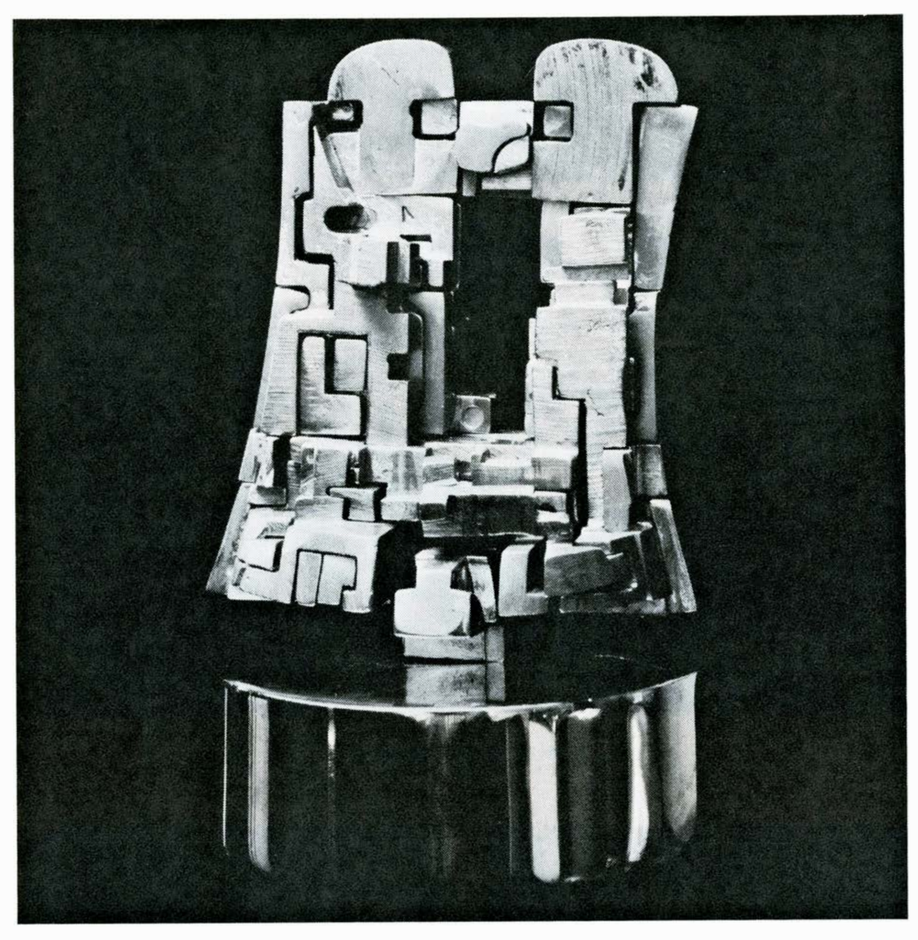


Mini María, *de Berrocal*





Un elemento interior de **Mini María (izquierda)** *y sus cinco piezas* **(derecha)**



Goliath *parcialmente desmontado*



Una escultura de Romeo y Julieta sobre otra escultura de su sepulcro

amantes. Dentro de los amantes hay una sorpresa que es preferible no describir en una revista de carácter familiar como ésta. El féretro es aún más sorprendente. Se descompone en 84 piezas que pueden recomponerse para dos equipos de mesa completos: 23 piezas de vajilla y cubertería, cuatro copas de vino, cuatro candelabros, un anillo de cena para hombre, otro anillo para mujer y un braserillo de mesa que es mayor que el ataúd.

Columbia Jet, mostrado en la ilustración superior de la página opuesta, fue encargado por las Líneas Aéreas Iberia, como regalo para sus ejecutivos. El pájaro es la paloma de la paz de Picasso. Su cuerpo es una jarra de agua que vierte el líquido por el pico al alzar un asa. El ornamentado pedestal es un vaso para beber. Desplegando las alas de la paloma, de las puntas de las alas descienden trenes de aterrizaje. Otra útil escultura de Berrocal es *Paloma Box* (La caja de Paloma), una escultura de alrededor de 30 centímetros de altura y anchura, que al abrirse resulta ser un joyero con 16 cajones forrados de terciopelo. Se puede alzar un espejo circular que, abierto, revela una escultura de una hija de Picasso, Paloma. Las piezas que componen la cabeza contienen, entre otras, dos brazaletes y dos cinturones. *Il Cavallo*, también de Berrocal, mostrada en la ilustración inferior de la página opuesta, tiene una clase de flexibilidad diferente. El caballo, compuesto por 14 piezas, tiene patas articuladas, con lo que puede adoptar diferentes actitudes.

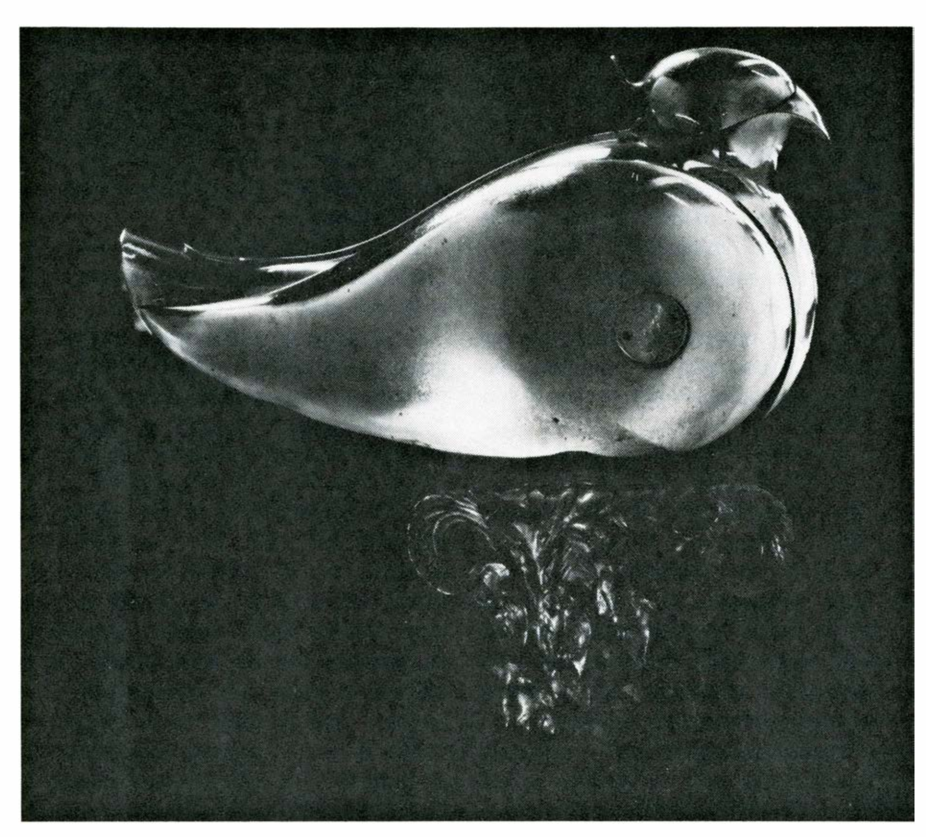
He descrito solamente una pequeña porción del trabajo de Berrocal. Su mayor escultura es su *Homage to Picasso* (Homenaje a Picasso), que mide más de 6 metros de largo y pesa 18 toneladas. Actualmente se encuentra en exhibición permanente en los Jardines de Picasso, en Málaga. Una versión miniatura en bronce, *Siextasis*, consta de 20 piezas enclavadas mediante pequeños imanes. Otra escultura de gran tamaño, tributo de Berrocal a su buen amigo Dalí, se encuentra en Madrid. Una versión miniatura en bronce, llamada *Delirium Tremens*, estará pronto disponible.

El agente exclusivo de Berrocal en los Estados Unidos es su viejo amigo y gran admirador James Smith Rudolf. Rudolf posee las dos librerías Centicore de Ann Arbor y la Centicore Arts, International (336 Maynard Street, Ann Arbor, Mich., 48108). Durante muchos años, Rudolf ha suministrado Berrocales a galerías y tiendas de arte y ha organizado alrededor de una docena de exposiciones individuales en distintas ciudades cada año. Los precios de las obras de Berrocal

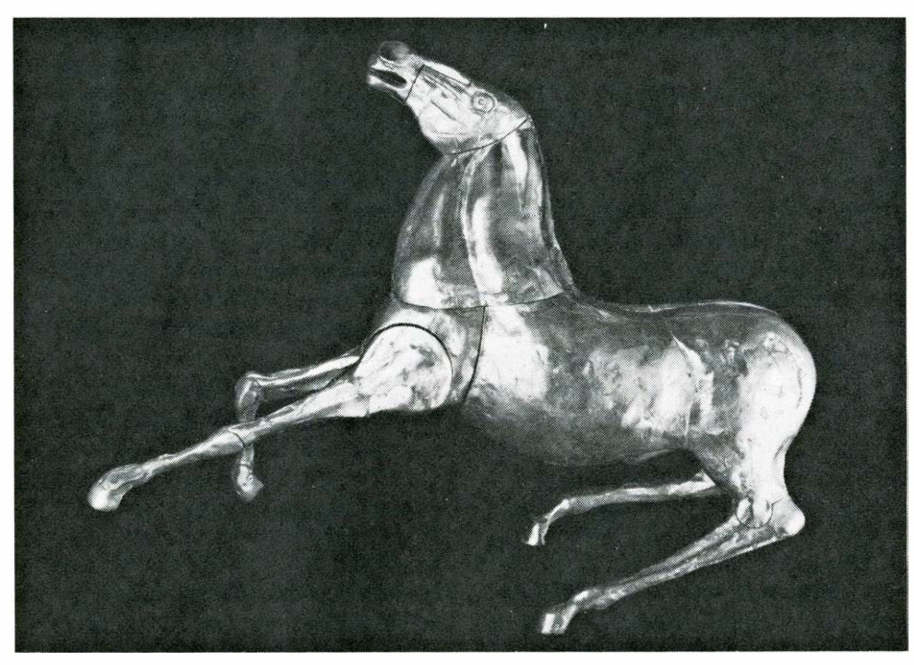
comienzan en unos 180 dólares para el micro (colgante), subiendo rápidamente a muchos miles de dólares para los múltiplos mini.

A excepción de los colgantes, no hay tornillos ni pasadores en ningún Berrocal. Los colgantes están suspendidos de una pieza maestra que se atornilla en él para evitar que pueda desprenderse accidentalmente de su cadena. Todos los múltiplos de oro y plata fina tienen piezas maestras que se insertan en su sitio con una llave ordinaria, de forma que la escultura no pueda desmontarse sin cooperación de su propietario.

Permítanme que anticipe una objeción. Podría preguntarse, ¿qué relación guarda el placer de desmontar una escultura co-



Columbia Jet, *de Berrocal*



Il Cavallo, *de Berrocal*

mo si fuera un rompecabezas, y de volverla a montar, con el arte? En cierto sentido, evidentemente, la respuesta es "nada"; pero también es verdad que eso no es todo. La belleza visual del arte siempre ha estado combinada de innumerables modos con otros valores: las emociones sexuales suscitadas por los desnudos; los sentimientos evocados por paisajes, marinas o retratos de familia; la función retórica del arte político o religioso; el valor didáctico de la ilustración de textos; el humor del arte cómico; la comodidad física procurada por sillas, camas y sofás diseñados para ser bellos; la utilidad de mesas, jarrones, botellas, copas, platos, cubiertos, coches, casas, edredones, barcos, relojes, herramientas, etc. El hallazgo, único, de Berrocal ha consistido en combinar los placeres visuales y táctiles con el gozo intelectual de resolver un rompecabezas mecánico. Si esta particular combinación no es de su agrado, los Berrocales no son para usted.

En la región de Quadling, en Oz (descrita en *The Emerald City of Oz)* la ciudad de Fuddlecumjig está habitada por una extravagante raza de gentes llamadas Fuddles (Atolondrados). Cada Fuddle está compuesto por centenares de piezas de formas fantásticas de madera pintada, que encajan entre sí, como un rompecabezas tridimensional. Cuando se les acerca un visitante, los Fuddles se descomponen en una cascada de piezas inconexas, para que el visitante tenga el placer de recomponerlo o recomponerla.

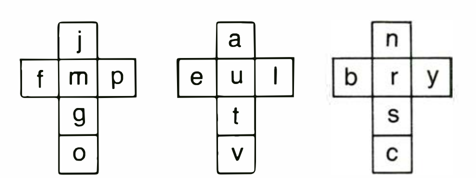
"Desde luego, son gente rara", dijo Em, la tía de Dorothy, cuando conoció a los Fuddles. "Pero la verdad es que no veo para qué sirven, si es que sirven de algo."

"Bueno, sirvieron para divertirnos algunas horas", respondió el Mago. "Yo creo que sí nos valieron de algo, después de todo."

"Me parece que son más divertidos que jugar a solitarios o al tres en raya", añadió el tío Henry. "Por mi parte, me alegro de haber visitado a los Fuddles."

He aquí las soluciones a los problemas numéricos del Doctor Matrix, propuestos en esta misma sección el mes pasado.

1. El problema consiste en demostrar que ningún número que comience por 9 y seguido por dígitos descendentes consecutivamente en orden cíclico (con o sin el 0) puede ser primo. A excepción del 2, todos los números primos han de terminar en 1, 3, 7 o 9, pues si la última cifra es par el número es divisible por 2, y si



Solución de los dados-calendarios del mes pasado

la última cifra es un 5, entonces el número es divisible por 5. Es fácil demostrar, sumando los dígitos y utilizando restos potenciales de 10, que si la sucesión termina en 1, 3, 5 o 7, el resto potencial correspondiente será múltiplo de 3, lo que demostraría que el número también es múltiplo de 3.

2. Al ordenar alfabéticamente los números de 0 a 1000, el penúltimo es 333, y el último es el 1.

3. Al ordenar alfabéticamente los números romanos de 1 a 1000, el último lugar está ocupado por XXXVIII, o sea 38.

4. En inglés, el mínimo número cuyo nombre contenga todas las vocales y la *"y"* es "one thousand twenty-five".

5. El siguiente término de la sucesión 10^3, 10^9, 10^{27}, 10^2, 10^0,... es $10^{0,60206}$..., es decir, 4. El primer término "one thousand" es el primer entero positivo cuyo nombre inglés contiene la letra *a;* el segundo término "one billion" es el mínimo número que contiene la *b;* el tercero, "one octillion" es el mínimo en contener la *c;* el cuarto "one hundred", es el primero en contener la *d;* el quinto, "one", es el primero que contiene la letra *e;* y el sexto, "four", es el primero que contiene la *f.*

6. Los cubos "desarrollados", que se muestran en la ilustración superior, aclaran cómo podrían rotularse letras minúsculas en las caras de tres cubos de forma que los cubos dispuestos en fila vayan dando las iniciales de los nombres ingleses de los meses del año. Nótese que esto es posible porque *u* y *n* son letras simétricas, y lo mismo ocurre con la *p* y la *d.* En realidad se pueden dejar dos caras sin rotular, pues la *b* puede utilizarse como *g* dándole la vuelta, y la *u* puede servir de *c* dándole un cuarto de vuelta.

Con toda la razón, muchos lectores han protestado de mi artículo del mes de diciembre, dedicado a la hipérbola, por decir que la trayectoria que describen los objetos lanzados al aire es una parábola. Tales caminos están muy cerca de ser parábolas, pero hablando estrictamente (y despreciando la resistencia del aire) un objeto lanzado al aire sigue una órbita elíptica alrededor del centro de gravedad de la tierra.

Taller y laboratorio

Más sobre los polarizadores y cómo usarlos en el estudio de la luz polarizada del cielo

Jearl Walker

Una de las fuentes naturales de luz polarizada es el cielo. Como escribió Lord Rayleigh, en 1871, dicha polarización es provocada por la dispersión de la luz solar al chocar con las moléculas del aire (fenómeno que también es causa del color azulado del cielo). En un día claro, gran parte de la luz que nos llega del firmamento viene linealmente polarizada, circunstancia de la cual se valen las abejas y otros animales para orientarse. Incluso hay razones para creer que los antiguos vikingos también se orientaban, cuando navegaban por mar abierto, gracias a la polarización de la luz del cielo.

Imaginemos un rayo de luz que, procedente del sol, se encuentre con una molécula de un gas de los que forman el aire. Consideremos como si la luz sin polarizar tuviera dos sentidos de polarización, un sentido contenido en el plano que determinan el sol, la molécula de gas y el observador situado en la tierra, y el otro sentido de polarización contenido en el plano perpendicular al anterior. Los dos sentidos de polarización (que llamaremos, respectivamente, paralelo y perpendicular) representan las direcciones a lo largo de las cuales oscilan los campos eléctricos que provocan la luz.

Ambos sentidos de oscilación dan lugar a que los electrones de la molécula con la que han chocado oscilen y reemitan luz. Sin embargo, la oscilación de un campo eléctrico que oscila a lo largo de una determinada dirección, al chocar con el electrón, no hace que éste emita radiaciones que oscilan a lo largo de dicha dirección, sino que el electrón suele emitir radiaciones que oscilan principalmente en un plano perpendicular al de la dirección de oscilación de la luz que lo originó. Progresivamente va emitiendo radiaciones, en menor proporción, en las direcciones que se van acercando a la dirección de la oscilación.

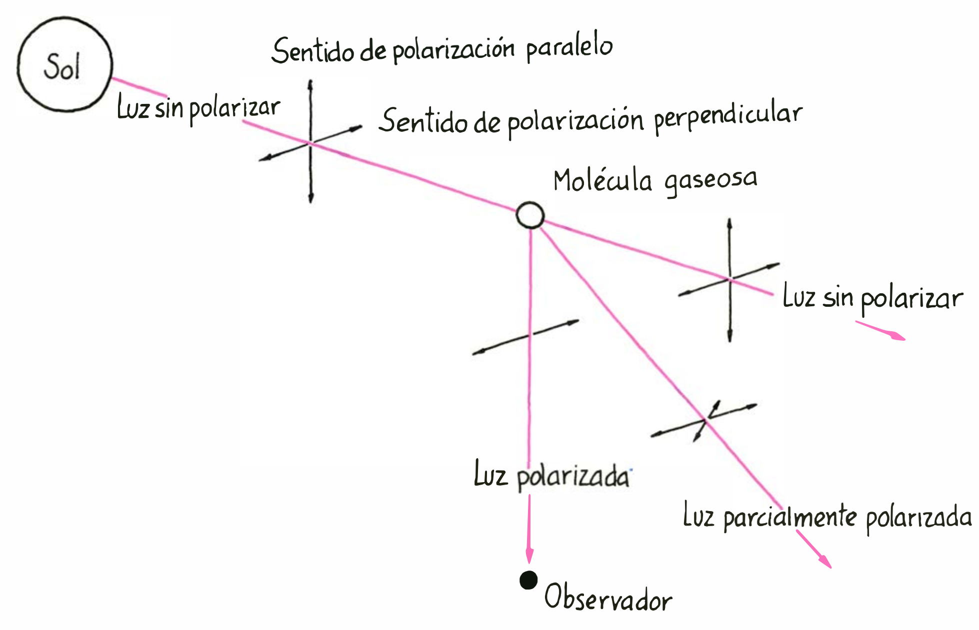
¿Cómo se explica con la hipótesis anterior la polarización de la luz solar dispersada a partir de una molécula de aire que el observador ve? Consideremos un observador situado directamente debajo de una molécula iluminada por la luz del sol cuando éste se encuentra en el ocaso. El sentido de polarización perpendicular de dicha luz al incidir en la molécula es desviado con bastante intensidad hacia el observador, mientras que el sentido de polarización paralelo no se dispersa en modo alguno. El resultado neto para el observador es que la luz procedente del sol sin polarizar, se polariza por la dispersión que sufre al chocar con las moléculas, y el observador lo que ve es luz polarizada.

El lector puede hacer un gráfico del tipo o grado de polarización de distintas zonas del cielo con un filtro polarizador. Primero oriente el filtro, es decir, averigüe en qué sentido polariza dicho filtro; para ello, examine la luz reflejada por la calzada o por cualquier otra superficie lisa (describí el sistema el mes pasado). Cuando el resplandor del reflejo en una superficie lisa queda bloqueado por el filtro, el sentido de polarización que éste deja pasar es el perpendicular a dicha superficie.

Ahora examine una zona del cielo. Dependiendo de la posición del sol, la luz que le llega de dicha zona estará polarizada principalmente en el sentido perpendicular al plano que contiene el sol, a la zona del cielo y al observador. Cuando el sol esté próximo al horizonte, la zona que da la luz con mayor grado de polarización será la que esté directamente encima de su cabeza. La polarización decaerá a medida que el observador vaya mirando a través del filtro las distintas zonas del cielo a lo largo de toda la dirección diametral en la que el sol se encuentre. (Nunca se debe mirar directamente al sol, ni siquiera a través de algo que el observador crea que es un filtro negro u obscuro. El filtro puede que sea eficaz en el rango de luz visible, pero puede permitir pasar cantidades peligrosas de radiaciones, bien del sector del infrarrojo bien del sector del ultravioleta. Tanto unas como otras pueden dañar los ojos antes de que el interesado se dé cuenta de ello.)

Cuando el sol esté bajo, la luz procedente de la zona del cielo cerca del horizonte, pero a 90 grados del sol (es decir, el Norte o el Sur geográfico cuando el sol se halle en el ocaso), estaría polarizada tan intensamente como la que viene de la zona del cielo justo encima de la cabeza del observador. Pero la luz que viene de aquel horizonte tiene que atravesar mucha zona de atmósfera terrestre, de manera que, probablemente, se ha dispersado más de una vez, debido a los choques con las múltiples partículas atomizadas o aerosoles existentes en la atmósfera. Todos estos hechos hacen decrecer el grado de polarización.

Las nubes destruyen también la polarización debido a las múltiples veces que



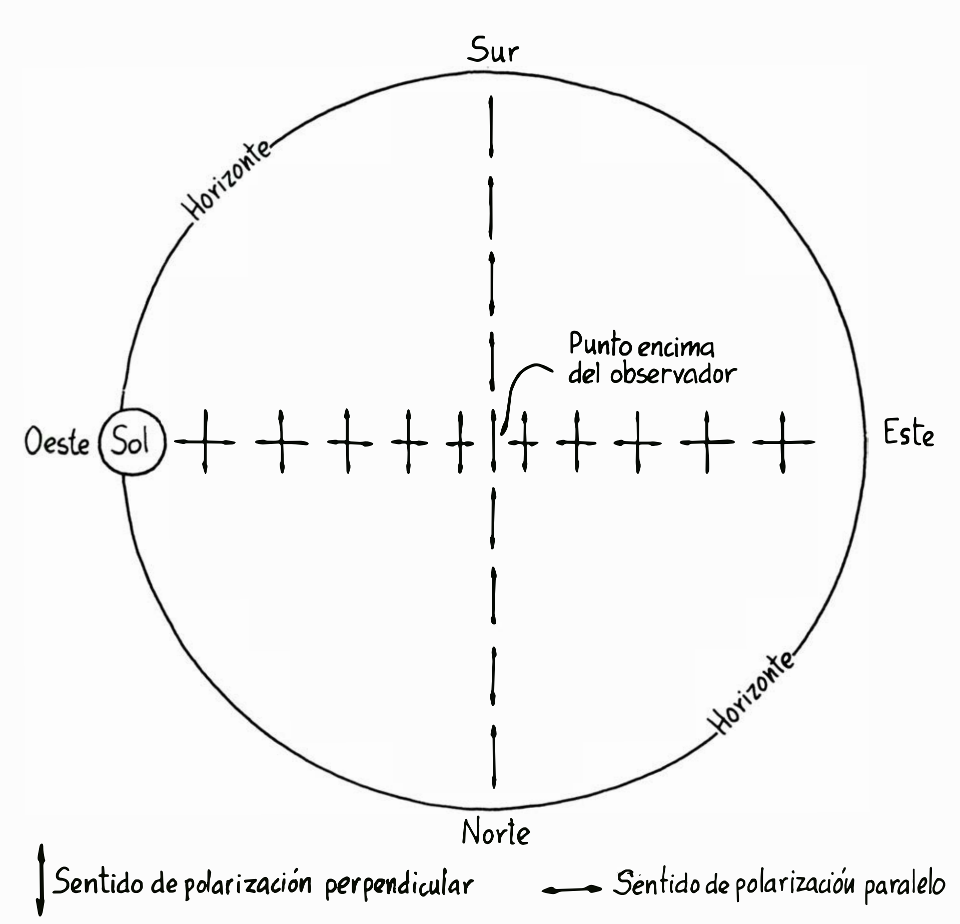
Polarización de la luz solar al ser dispersada por las moléculas gaseosas del aire

la luz se dispersa al atravesarlas; este efecto es utilizado con bastante frecuencia por los fotógrafos. Con un filtro polarizador orientado de forma adecuada y colocado delante del objetivo de una cámara fotográfica, la luz polarizada que viene del fondo del cielo es bloqueada en una proporción bastante grande (función de su grado de polarización), mientras que la luz que viene de las nubes, aunque también en parte esté bloqueada por el filtro, lo es en bastante menor proporción, puesto que su grado de polarización es bastante menor también. A consecuencia de lo cual, en la foto las nubes destacan sobre el fondo del cielo.

Se puede construir un aparato muy útil para determinar la polarización de la luz del cielo a partir de varios filtros polarizadores pequeños o a partir de una lámina polarizadora grande. En cualquiera de los dos casos, se deben recortar doce triángulos isósceles de manera que el ángulo distinto sea de 30 grados y el sentido de polarización de cada uno de los mismos sea paralelo a la base. Estos triángulos se colocan sobre una superficie de cristal transparente situándolos de manera que formen un dodecágono. Los bordes del triángulo se deben sujetar al cristal o bien con cinta de celofán o con pegamento. Si se usa este segundo procedimiento hay que tener cuidado de no pegar la superficie completa de los triángulos al cristal, ya que el medio adhesivo puede alterar la polarización de la luz que filtre el conjunto.

Cuando el observador mire al cielo a través del dodecágono, verá triángulos iluminados y otros obscuros, de acuerdo con la polarización de la luz del cielo. Ahora ya no es necesario hacer girar el sistema como cuando se usaba sólo un filtro polarizador. Para hacer que la polarización, o al menos los colores se destaquen más, basta colocar una o dos capas de cinta de celofán transparente sobre los triángulos antes de recortarlos. La cinta se debe orientar de manera que forme 45 grados con el sentido de polarización de los filtros de forma que cada triángulo pueda actuar de hecho como una lámina de media onda del tipo de las que describimos el pasado mes. La cinta debe colocarse en la cara de los triángulos que mira al cielo. La luz linealmente polarizada del cielo atraviesa la cinta y sufre un desfase entre los ejes rápido y lento de la cinta. La luz sale con una polarización diferente de acuerdo con su color.

El filtro sobre el cual la cinta se pega determina qué color se verá. A través de este montaje, equivalente a una lámina de media onda, el cielo aparece de diferentes colores por los diferentes triángulos. Con una sola capa de cinta adhesiva, encontré una variedad de colores en todo el círculo del aparato, que iba desde el azul claro (que es el que tiene el cielo realmente) al intenso. Con dos capas, la variación va desde el amarilloverde al azul.



Esquema del cielo donde se indica el sentido de polarización de la luz según el modelo de dispersión de Rayleigh

Hoy los filtros polarizadores pueden emplearse para determinar la polarización del cielo con el fin de orientarse en la navegación. Hasta 1928 no fue desarrollado –por Edwin H. Land– un polarizador de este tipo. Anteriormente, la polarización del cielo sólo se podía comprobar con los cepillos de Haidinger (curiosa figura que algunas personas ven cuando miran la luz polarizada) o con los cristales dicroicos y birrefringentes que describí el mes pasado.

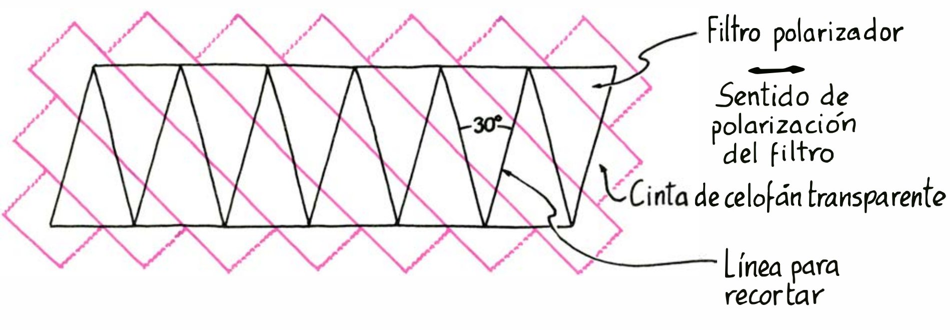
El artilugio que permitió a los vikingos cruzar el Atlántico con la ayuda de la luz linealmente polarizada del cielo pudo haber sido muy bien un cristal dicroico. No teniendo brújula, podían perderse cuando el sol se ocultaba por el horizonte, cosa bastante frecuente por las altas latitudes de su región de origen. Varios tipos de cristal pudieron haberles servido como "piedras del sol", pero el arqueólogo danés Thorkild Ramskou cree que probablemente usaron un simple cristal de cordierita.

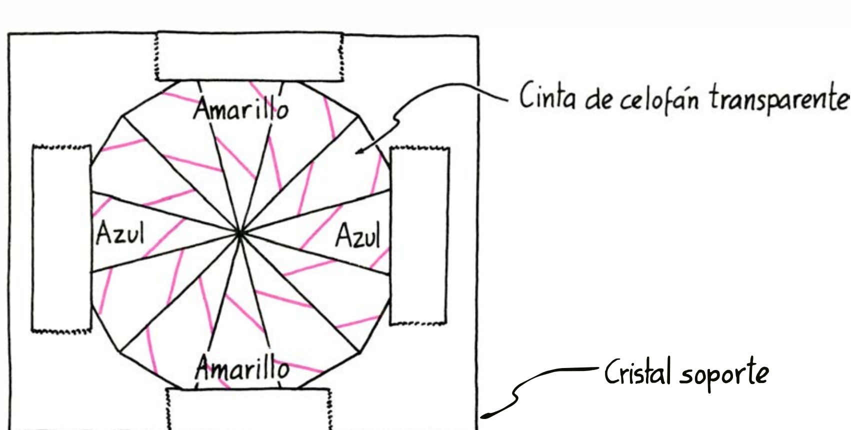
A Ramskou le dio la idea un muchacho de diez años, cuyo padre era piloto de la compañía aérea Scandinavian Airlines. Este chico encontró relación entre las viejas historias de las piedras de sol y el instrumento llamado brújula del crepúsculo, con el que su padre localizaba el sol incluso cuando estaba oculto por las nubes o había descendido tras el horizonte. Un piloto moderno detecta, con un filtro polarizador, el sentido de polarización de la luz procedente de una parte despejada del cielo; el instrumento señala entonces una dirección perpendicular a aquel sentido y, por tanto, lo que indica es la dirección diametral en la que el sol se encuentra. Usando la brújula del crepúsculo se puede localizar la posición del sol con bastante precisión, conociendo aproximadamente en qué dirección se encuentra el mismo.

La cordierita se comporta dicroicamente cuando sus ejes rápido y lento son perpendiculares a la luz incidente que le llega del cielo. Rotando el cristal, pero manteniéndolo perpendicular a la dirección de la luz, se provocan una serie de

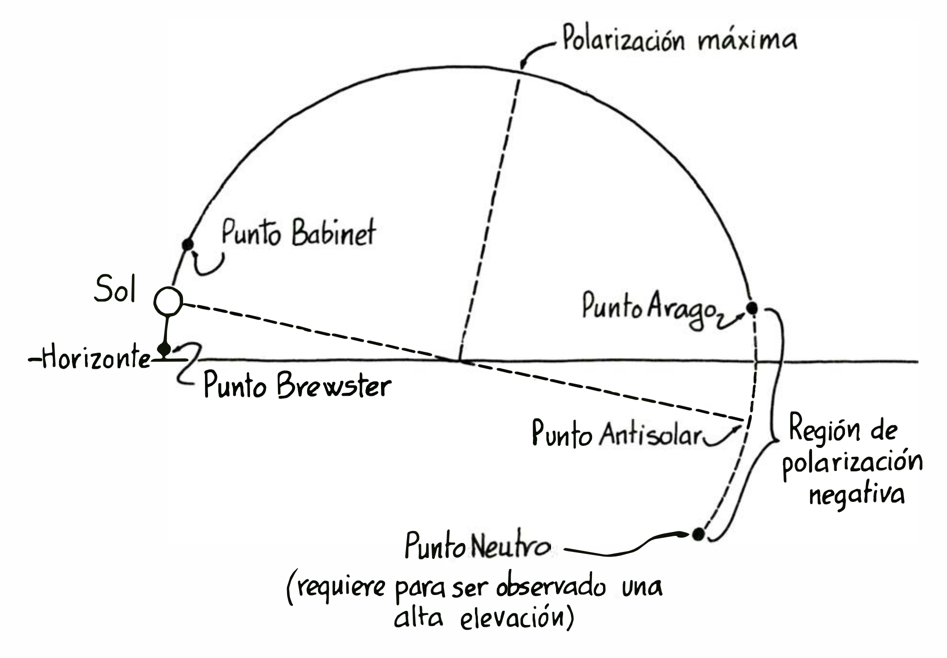
ejes distintos paralelos al sentido de polarización de la luz incidente. Con el eje rápido alineado, el cristal es o bien claro o ligeramente amarillo, al mirar a su través. Si es el eje lento el que está alineado, se ve azul. Un piloto vikingo cayó en la cuenta de cómo había que orientar (con respecto al sol) su cristal para que diera luz azul, por tanto, él podría encontrar el sol aunque estuviese escondido tras el horizonte u oculto detrás de algunas nubes, sin más esfuerzo que reorientar su cristal, enfocando una parte despejada del cielo, hasta que el cristal diera de nuevo el color azul (si todo el cielo estaba cubierto por las nubes, no había, por tanto, ninguna luz polarizada y el vikingo ese día no estaba de suerte).

Ramskou voló junto con el piloto de la Scandinavian Airlines para comparar la indicación de la localización del sol hecha con la cordierita y con la brújula del crepúsculo. Así comprobó que la indicación sobre dónde estaba el sol hecha con su "piedra del sol" tenía un error de aproximadamente 3 grados. También detectaba la dirección en que se encontraba el sol, incluso hasta cuando éste se hallaba 7 grados por debajo del horizonte. Si el lector tiene la posibilidad de disponer de un cristal de cordierita o de algún otro cristal dicroico similar como la turmalina, puede tratar de hacer un experimento parecido.





Forma de hacer un dodecágono de láminas de media onda para estudiar la polarización de la luz del cielo



Puntos neutros y zona de polarización "negativa"

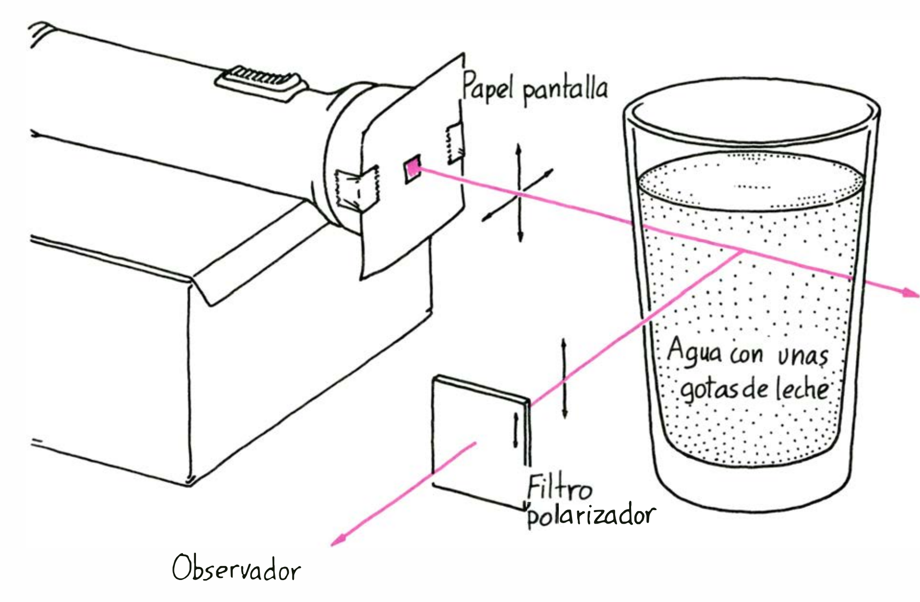
Gran parte de lo que se sabe acerca de la luz polarizada del cielo está basado en los trabajos de Lord Rayleigh. Su tesis sobre la polarización del cielo se puede observar de la siguiente forma. Primero, si el aire está lo suficientemente poluto, la dispersión múltiple de la luz, así como la dispersión provocada por partículas de tamaño aproximado (o incluso mayor) a la longitud de onda de la luz visible, destruye el sentido de polarización en el modelo de Rayleigh. Puede verse esta destrucción en días de gran polución atmosférica, ya sea causada por el hombre o por motivos naturales, como, por ejemplo, las erupciones volcánicas.

La otra aplicación del modelo de Rayleigh es más curiosa. Si se examina atentamente el cielo cerca del sol o cerca del punto antisolar (que es el punto exactamente opuesto al anterior en el espacio), encontraremos varios puntos que están sin polarizar y también una pequeña zona con el sentido de polarización opuesto, denominada polarización negativa. A los puntos se les llama puntos neutros y se le han dado diversos apellidos según los varios descubridores que ha tenido. La justificación de los puntos neutros y de la zona de polarización negativa es todavía en ciertos aspectos motivo de controversia, pero parece ser debido tanto a la dispersión múltiple de la luz solar como a la dispersión debida a los aerosoles, principalmente en las zonas bajas de la atmósfera.

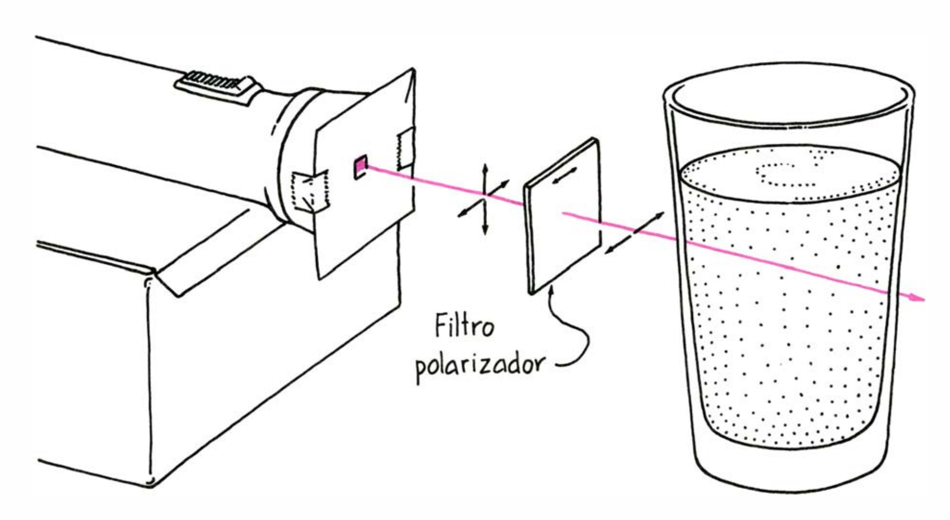
El punto neutral que separa el área de polarización negativa del área de polarización positiva (normal) es conocido como "punto de Arago" (en honor de D. F. J. Arago, astrónomo y físico francés de principios del siglo XIX). Yo puedo localizar el punto de Arago esperando hasta que el sol esté cerca del horizonte, y entonces busco la región de polarización negativa cerca del punto antisolar. La posición del punto de Arago (y, por tanto, la zona angular de la región de polarización negativa) depende de la longitud de onda de luz que se considere, de la elevación del sol y de la amplitud de la reflexión de la luz en la superficie de la tierra. Por ejemplo, si se coloca un

filtro azul o rojo sobre el filtro polarizador (se puede usar para ello un papel de celofán azul o rojo), verá que el punto de Arago para el azul está situado más arriba que para el rojo. La posición de dicho punto también depende de la elevación en que se encuentre el sol; mientras más bajo esté menor será la zona de atmósfera cercana a la tierra que esté iluminada.

Si el observador empieza con el punto de Arago mirando al punto antisolar con el sol en el ocaso a su espalda, al principio el punto de Arago se mueve ligeramente hacia fuera (arriba) desde el punto antisolar y retrocediendo como si buscara al sol en el horizonte. En el ocaso, la atmósfera que está más cerca del suelo, en donde la dispersión múltiple y la dispersión aerosolar responsable de la polarización negativa tienen lugar, no está demasiado bien iluminada, cuando, por el contrario, la parte superior de la atmósfera aún lo está. En consecuencia, la región de polarización negativa se hace más pequeña y la región de la polarización positiva aumenta, moviendo el punto de Arago hasta acercarlo al punto antisolar. Con el sol por debajo del horizonte, el punto de Arago se mueve otra vez hacia fuera del punto antisolar. El movimiento total del punto de Arago con respecto al punto antisolar es de 5 grados aproximadamente. El tercer efecto sobre la posición del punto de Arago es atribuible a la luz reflejada por la superficie terrestre. Con una gran reflexión, como en el caso de que se reflejara sobre una gran masa de agua, el punto desciende unos cuantos grados.



Efecto de dispersión producido por una suspensión coloidal de leche en agua



Efecto de una dispersión cuando la luz incidente está horizontalmente polarizada

Mientras que la luz dispersada por las moléculas puede ser fuertemente polarizada y la dispersada por partículas grandes puede ser no polarizada o polarizada solo débilmente, las partículas de niebla de tamaño intermedio producen también un grado intermedio de polarización. Esta polarización intermedia tiene efectos interesantes en los objetos distantes obscurecidos por la niebla. Por ejemplo, las montañas de San Gabriel, en California, se ocultan con frecuencia debido a la nube de polución que se forma sobre Los Angeles. Sin embargo, usando gafas de sol polarizadas o cualquier otro filtro polarizador, se pueden ver estas montañas a pesar de la polución. Warren H. White, del Meteorology Research Inc., explica que la diferencia de visibilidad se debe al grado de polarización de la nube de polución, lo cual permite determinar el tamaño aproximado de las partículas de ésta. Orientando el filtro de una manera determinada se distinguen las cumbres de las montañas recortadas sobre el firmamento, pero los accidentes de las laderas no son visibles. Sin embargo, si se gira el filtro 90 grados cambia el sentido de la polarización transmitida, lo que hace que desaparezcan de la vista las cumbres de las montañas y que sean perceptibles los accidentes del relieve.

Si las partículas de la nube de polución fueran muy pequeñas comparadas con la longitud de onda de la luz visible, supongamos que tuvieran diámetros inferiores a 0,05 micrones, polarizarían la luz como lo hacen las moléculas individuales en el modelo de Rayleigh. Si las partículas fueran mayores que las longitudes de onda de la luz visible, con diámetros superiores a los 5 micrones, la luz dispersada se polarizaría muy débilmente. Supongamos un rayo de luz solar que llega a las montañas formando un ángulo de 45 grados con la vertical. Con la primera orientación del filtro se distinguen las cimas de las montañas porque la luz del cielo sobre el que se recorta está más polarizada que la dispersada por la nube de polución situada entre las cimas y el observador. La luz del cielo es, por tanto, más brillante que la que reflejan las cimas, y el contraste es evidente.

Al girar el filtro 90 grados, la luz del cielo y la de las cumbres son bloqueadas, pero se hacen visibles algunos accidentes de las laderas. Son las zonas orientadas adecuadamente para polarizar la luz solar reflejada, de forma similar a lo que ocurre en la superficie de una carretera. Aunque esta luz está polarizada sólo débilmente, se distingue con facilidad porque la luz del resto de la montaña está polarizada por la nube de polución de tal manera que se bloquea con la nueva orientación del filtro. Debido a que la nube de polución polariza la luz más intensamente que los accidentes de la superficie, pero no tan intensamente como el aire limpio más lejano, podemos deducir que los diámetros de las partículas de polución deben medir entre 0,05 y 5

micrones. Esta conclusión ha sido confirmada por investigaciones más profundas que sitúan el rango entre 0,1 y 1 micrón.

He visto un efecto similar al observar cómo desaparecen los barcos debido a la nube de polución de Cleveland cuando navegan hacia el Lago Erie. Si el lector vive en un entorno que tenga un nivel comparable de polución, quizá desee repetir el experimento de White y buscar la correspondencia entre el tamaño del aerosol y la estación. El mismo experimento se podría llevar a cabo con las nieblas azules que cubren algunos tipos de montañas, tales como las de Blue Ridge en el oeste de Virginia y las Montañas Azules en Australia. Se cree que estas nieblas tienen su origen en las pequeñas partículas de cera que desprende la vegetación como consecuencia de los campos electrostáticos relativamente altos que existen cerca de los extremos del follaje. Las partículas son inferiores a los 6 micrones de diámetro y por ello producen cierta polarización al dispersar la luz del sol.

La polarización parcial de la luz dispersada por un aerosol puede ser simulada en la cocina por medio de un experimento simple. Tape casi todo el frente de un foco (o cualquier fuente de luz similar) y deje que pase sólo un rectángulo estrecho de luz. Coloque el foco de manera que la luz pase a través de un vaso de agua limpia, apague las luces de la habitación y examine el cristal del vaso por el lado perpendicular al haz de luz (es mejor un haz de luz no muy potente porque así se evita la iluminación excesiva de las paredes del vaso).

Si el agua está muy limpia, será demasiado transparente para dispersar luz perpendicularmente y el vaso aparecerá obscuro. Cada molécula de agua está dispersando luz en todas direcciones en la habitación, pero la luz dispersada por una molécula interfiere anulando la luz dispersada por otra en todas direcciones, excepto la que va en la dirección del haz de luz. El rayo de luz se reconstruye, pues, en esa dirección y no existe dispersión de luz neta en otras direcciones.

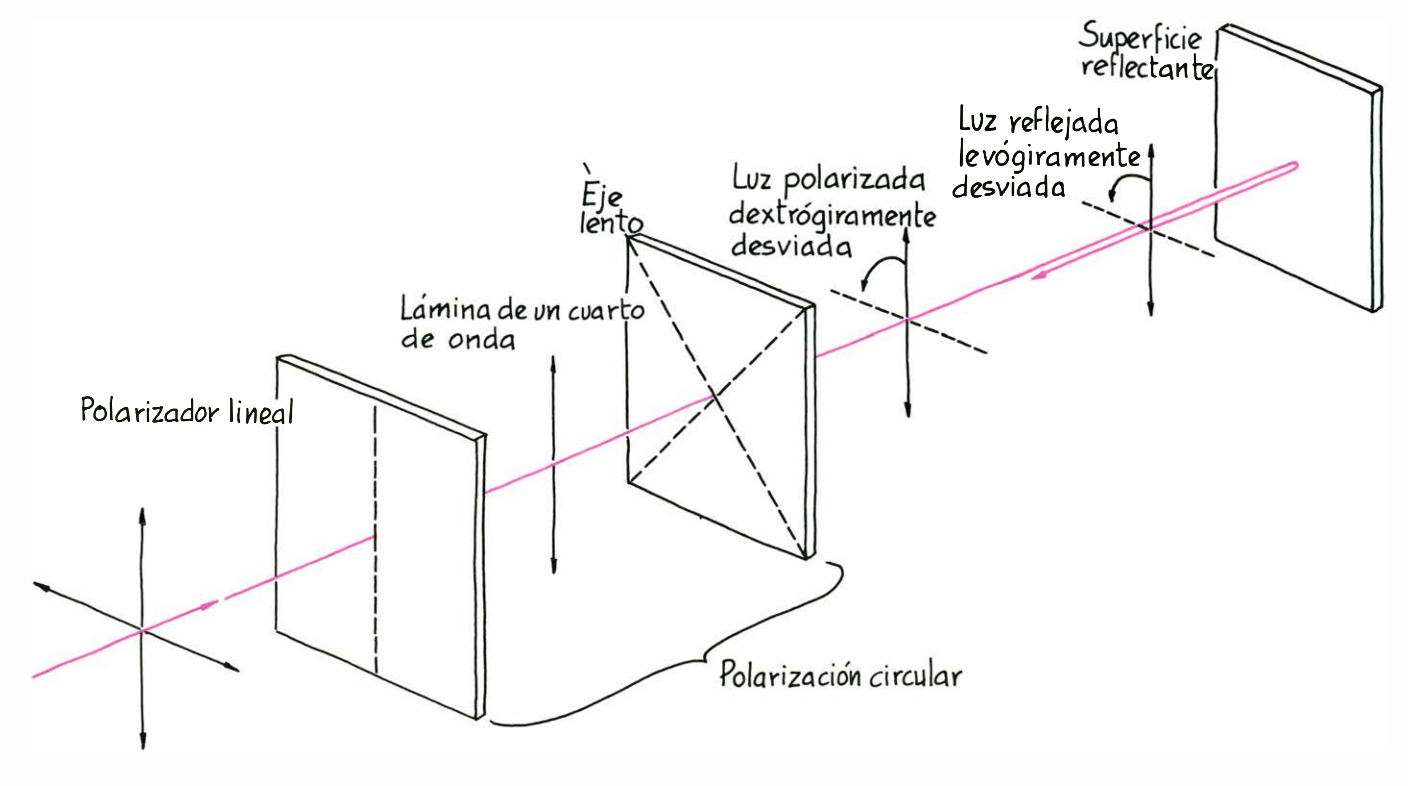
Para continuar con el experimento añada unas cuantas gotas de leche. Antes de comprobar la dispersión perpendicular, observe los colores de la leche. El azul es dispersado en línea recta más fuertemente que el rojo y los colores entre azul y rojo del espectro se difunden en ángulos intermedios. Si se mira la leche desde el lado de donde está iluminado el foco, la leche parece azulada. Si se mira desde el otro lado (hacia el haz de luz) la misma aparece rojiza.

Esta distribución de los colores es similar al aspecto general del cielo durante el día, aunque la dispersión no es exactamente la misma. Cuando la luz es dispersada por moléculas de aire según el modelo de Rayleigh, la luz azul se dispersa más fuertemente que la roja. La consecuencia es que todo el cielo es azul la mayor parte del día. Cuando el sol se está poniendo o saliendo, el azul es dispersado debido a la luz próxima al horizonte, quedando fundamentalmente el rojo, que es la razón por la cual una puesta o una salida de sol son rojas. La luz difundida por la leche no es dispersada por moléculas individuales porque la leche no está disuelta en el agua. La leche forma más bien una suspensión coloidal de pequeños glóbulos.

Esta suspensión dispersa luz en ángulos diversos con respecto a la dirección del foco de luz y por ello polariza la luz dispersada. Con un filtro polarizador, el lector puede comprobar la polarización de la luz dispersada perpendicularmente al haz de luz. Si el haz de luz es horizontal, el filtro transmitirá luz dispersada con un sentido de polarización vertical. A medida que se van añadiendo más gotas de leche disminuye el grado de polarización en la luz dispersada.

Cuando el agua con leche es claramente blanca y desaparecen el rojo y el azul, la luz dispersada ya no se polariza. La densidad de los glóbulos de leche suspendidos es en esos momentos tan alta que la mayor parte de la luz dispersada lo ha sido de forma múltiple cuando llega al ojo. Lo mismo que ocurre con la dispersión atmosférica, la dispersión múltiple destruye la polarización.

Yo no he encontrado diferencia en los resultados obtenidos usando leche des-



Absorción de luz reflejada por un polarizador circular

cremada y leche sin descremar, a pesar de su diferente contenido en grasa. El lector quizá quiera probar con otras suspensiones coloidales, o con azúcar, para ver la diferencia entre una solución (en la cual la luz es dispersada por moléculas individuales) y una suspensión coloidal. Con unas cuantas gotas de leche en el agua se puede también demostrar un resultado del que he hablado anteriormente: la luz no se dispersa a lo largo de los ejes de polarización de la luz incidente. Esto se puede comprobar fijando un filtro polarizador al foco y examinando con otro filtro la luz dispersada perpendicularmente al haz de luz. Si el haz de luz del foco se polariza verticalmente, el ojo está recibiendo luz dispersada polarizada. Si el haz de luz es polarizado horizontalmente, el ojo está recibiendo luz no dispersada porque permanece en los ejes de polarización de la luz incidente.

El mes pasado vimos cómo se puede producir luz polarizada circularmente; se pasa la luz linealmente polarizada a través de una lámina de cuarto de onda colocada de tal forma que los ejes de la polarización estén a igual distancia entre el eje rápido de la lámina y el lento. En la longitud de onda para la cual la lámina es precisamente de un cuarto de onda aparece la luz polarizada circularmente. El sentido de rotación del vector de polarización que emerge depende de la orientación de la polarización incidente y de los dos ejes de la lámina. Convencionalmente se llama a la rotación "dextrógira" si tiene el sentido de las manecillas de un reloj según se mira hacia la fuente de luz y "levógira" si la rotación tiene el sentido contrario.

Si un polarizador circular concreto produce, por ejemplo, luz dextrógira, ésta puede también pasar a través del montaje en la dirección opuesta, aunque la luz emitida, finalmente, queda polarizada linealmente por el filtro. Si una luz polarizada levógira pasa a través de un polarizador circular dextrógiro, una vez que sale de la lámina de cuarto de onda, la luz es absorbida. Este efecto ha sido explotado para reducir el reflejo en las pantallas de televisión, osciloscopios, ventanas de aviones y en aquellos sitios donde el reflejo es molesto.

Así, se puede colocar delante de la pantalla de un televisor una placa polarizadora circular que se produce comercialmente. La placa es un conjunto de filtros polarizadores lineales pegados a 45 grados entre el eje rápido y lento de una lámina de cuarto de onda, con el filtro lineal en el exterior. La luz de la habitación que pasa a través de la lámina se polariza en el sentido dextrógiro; por ejemplo, cuando la luz se refleja en la pantalla de televisión, la reflexión convierte el sentido de polarización de la luz en luz levógira, pero la hoja absorbe ese sentido. Así, la luz de la habitación que de otra forma se habría reflejado en la pantalla interfiriendo la visión, es absorbida por la hoja circularmente polarizada.

Para demostrar este tipo de absorción, el lector puede construir un polarizador circular a partir de una lámina de cuarto de onda que describí el mes pasado. Coloque el polarizador en una moneda pequeña o cualquier otro objeto brillante. Si el polarizador lineal, que es parte del polarizador circular, está en el fondo del polarizador, la moneda brilla cuando es observada a través del montaje. Si el polarizador lineal está encima, la moneda es obscura, debido a la absorción de la luz reflejada. Si se quita el polarizador de encima de la moneda, ésta vuelve a ser brillante a medida que le da la luz sin tener que pasar a través del polarizador.

Si se cambia la moneda por una lámina lisa de aluminio y se repite el experimento, llegaremos a los mismos resultados. Podemos posteriormente hacerle a la lámina un pliegue en forma de *V* y poner el polarizador circular sobre la *V* con el filtro lineal hacia arriba. La parte lisa de la lámina se verá obscura otra vez, pero el pliegue estará brillante porque está actuando como un doble espejo.

Supongamos que un polarizador circular produce luz polarizada dextrógira, el componente de la luz que choque con un lado del pliegue se reflejará al otro lado antes de volver a aparecer, cambiando así su polarización de dextrógira a levógira; después volverá a ser dextrógira antes de llegar al polarizador otra vez. Su regreso a luz polarizada dextrógira quiere decir que pasa a través del polarizador circular sin absorción, mostrando, por tanto, una *V* brillante.

El lector puede continuar haciendo investigaciones con su polarizador circular haciendo uso de las láminas de cuarto y media onda. Si la luz polarizada circular levógira pasa a través de una lámina de media onda, el observador podrá ver luz polarizada dextrógira. Si es, en cambio, pasada a través de una lámina de cuarto de onda, verá luz polarizada linealmente. Quizá le interese repetir la demostración de absorción con un reflector colocando una lámina de media onda entre el polarizador circular y el reflector. ¿Absorbe todavía el polarizador circular la luz reflejada?

Libros

A propósito de los nuevos instrumentos musicales electrónicos e introducción a la genética humana

A. Lewin-Richter y A. Sánchez Cascos

NUEVA GENERACION DE INSTRUMENTOS MUSICALES ELECTRONICOS, por Juan Bermúdez. Marcombo, S.A., Boixareu Editores 1977.

En 1978 celebramos el treinta aniversario de una nueva técnica de composición musical, que si bien hoy recibe el nombre de *música electroacústica,* o todavía, *música electrónica,* Pierre Schaeffer, su iniciador, la denominó *música concreta.* En 1948 se realizaron en París los primeros ensayos con tocadiscos y un material grabado a través de micrófonos. El progreso ha sido vertiginoso: la aparición de los magnetofonos comerciales de cinta, primero con soporte de papel, y después con soporte de plástico, crearon un nuevo vehículo de almacenamiento de material sonoro. En 1951 entró en juego la escuela de música electrónica de Colonia, que utilizaba un material de origen puramente electrónico, siguiendo unas teorías de composición musical muy puristas derivadas del serialismo de Webern. La pugna ideológica entre las tendencias de París y Colonia llevaron a una especie de vía intermedia, cuyo máximo exponente es, sin duda, "El Canto de los Adolescentes", de Stockhausen, aceptándose entonces el concepto de música electrónica para ambas tendencias, que se aúnan, finalmente, hacia 1960. Gracias al esfuerzo de los centros de París y Colonia y al impulso del movimiento paralelo de la *tape music* (música para cinta) de Vladimir Ussachevsky, de la Universidad de Columbia, Nueva York (iniciado en 1951), se crean laboratorios de música electrónica en todo el mundo, con una máxima proliferación en los Estados Unidos. En España existen en este momento cuatro: Alea en Madrid (Luis de Pablo), Phonos en Barcelona (José M. Mestres-Quadreny y Andrés Lewin-Richter), Actum en Valencia (José Luis Berenguer) y San Sebastián (José Luis Isasa). Estos laboratorios o estudios de música electrónica constituyen centros de formación e investigación musical, con producciones para los más variados campos de aplicación: música electroacústica pura, música para instrumentos y cinta, música aplicada para cine, teatro y ballet.

La multiplicación de los estudios de música electrónica, especialmente en los Estados Unidos, tanto en el campo de los departamentos de música de las universidades, como en el campo privado, y, posteriormente, en el área interpretativa rock, gracias al apostolado de Vladimir Ussachevsky, impulsaron la creación de sintetizadores de tipo comercial basados en las necesidades más apremiantes de los laboratorios de música electrónica.

En realidad, el primer sintetizador fue construido por RCA, conocido por tipo Mark I (1951), con el objeto de sintetizar sonidos instrumentales, a imitación de los existentes; se perfeccionó, en 1958, con el Mark II, que poco después se abandonaría debido a su escasa utilidad. El compositor Milton Babbitt se interesó por él y consiguió los fondos necesarios para trasladarlo y ponerlo a punto en los laboratorios de la Universidad de Columbia, utilizándose desde 1960 con fines compositivos dentro del contexto del Columbia Princeton Electronic Music Center.

Hacia 1966 empezaron a aparecer las primeras unidades modulares construidas por Buchla Associates, de Berkeley, California y R.A. Moog Co., de Trumansburg, Nueva York, que poco a poco tomaron la forma de unidades compactas y portátiles o semiportátiles, utilizando el nuevo concepto de *control de tensión,* gracias a la técnica de los amplificadores operacionales. Este es el tema central del libro que nos ocupa.

En el primer capítulo nos habla de la evolución tecnológica de los métodos de generación de sonido, muy pobremente, sin el fondo musical requerido y con unas ilustraciones harto banales. El segundo se dedica a la física del sonido, de forma divulgadora, muy somera y teniendo poco claros los conceptos musicales. En la "Producción electrónica del sonido" entra de lleno en un terreno más firme y con buenos conocimientos en la materia de la técnica operacional.

Dedica un capítulo muy extenso al sintetizador, su organización y su constitución: el teclado, su estructura electrónica y su problemática, los osciladores controlados por tensión (VCO), que admiten señales externas que complementan o modifican las propias; los filtros, los amplificadores controlados por tensión (VCA), los generadores de ruido (blanco, en especial), los moduladores en anillo, la reverberación, los generadores de envolvente, los mezcladores, los osciladores de modulación, los secuenciadores, los dispositivos "Sample and Hold", los seguidores de envolvente ("amplitude follower"). Todo ello se explica con detalle, utilizando unos textos en inglés (no siempre traducidos con claridad), si bien comprensibles por quien domine el tema.

El capítulo quinto es más complejo, pues hace unos considerandos para la construcción de módulos, de acuerdo con las necesidades y medios de cada estudio de música electrónica; trata cuestiones relativas a los bancos de osciladores, las fuentes de alimentación y su problemática y los esquemas variantes de los módulos citados en el capítulo anterior. Lo que sorprende es el último cuadro de este capítulo, demostrativo de una falta de originalidad, evidenciando, a mi entender, que todo el capítulo es copia de otro publicado en los Estados Unidos, y el autor no se ha preocupado por adaptarlo a España o a Europa, y, por ciertos errores que se escapan, dan idea de que no se ha sabido interpretar el material de origen.

El capítulo sexto se dedica a tres fabricantes de sintetizadores comerciales, hoy disponibles en el mercado. Moog ya no es independiente, como en la década de los 60; hoy pertenece al grupo Wurlitzer, fabricante de órganos electrónicos, toda vez que los sintetizadores han encontrado un campo comercial muy amplio como complemento de los órganos electrónicos y como instrumento de los grupos de música pop y rock, por la cantidad de efectos espectaculares que con ellos se puede obtener, como instrumentos propiamente dichos y como transformadores de los sonidos de otros instrumentos. Buchla dejó de existir en el mercado. En su lugar está ahora Arp. El modelo más extendido en España es, probablemente, el producido por Electronic Music Studios (E.M.S.) de Londres, ideado por Peter Zinovieff, quien ha creado unas unidades compactas, reducidas, muy manejables, muy versátiles, que se constituyen perfectamente en núcleo de un estudio de música electrónica y tiene grandes posibilidades de utilización en vivo.

Se dedica el capítulo séptimo a la descripción de elementos especiales, verbigracia, los desplazadores de frecuencia que parten del primitivo Klangumwandler (transformador de sonido) del alemán Harald Bode, inicialmente construido para el estudio de Baden Baden (1955); éste, mediante realimentación, puede crear automáticamente escalas ascendentes y descendentes. También se detiene en los moduladores de fase, que originan efectos de retardo, en los controladores de cinta tomados de las ondas Martenot, con el objeto de crear glisandos, en los convertidores de tono a tensión y en los controladores entre dos y tres ejes, tipo joystick. Hace una breve referencia a un estudio mixto con ordenador y sintetizador, campo que está en sus comienzos, a propósito del cual el autor no se rompe la cabeza. Sí añade unas notas sobre el mellotrón y el controlador de percusión. Echamos de menos la presentación de un filtro de tercio de octava, que, por su coste, siempre se deja de lado, y una mención al reverberador AKG. Por otra parte estimamos de mal gusto muchas de las ilustraciones utilizadas, por banales, superfluas y poco aclaratorias del texto.

El capítulo octavo resulta un verdadero cajón de sastre: programación del sintetizador, discografía y esquemas programados. Es lástima que la discografía sea reducida, sin explicación ni clasificación, y tendenciosa, con un claro énfasis en la música popular.

Todo ello nos lleva a las siguientes conclusiones. La intención del autor es óptima: aportar unos conocimientos técnicos a músicos, especialmente a aquel sector de músicos pop y rock que hacen uso musical de los sintetizadores, sin comprender su fondo técnico, haciéndoles vislumbrar un mundo de posibilidades tímbricas e interpretativas. Pongamos también, entre los aciertos, su labor de recopilación de esquemas. Sin embargo, la pobreza del texto es notable, pues se limita a recopilar una serie de manuales y artículos de revistas, los traduce (no sin errores, defectos, definiciones insuficientes y desconocimiento del lenguaje musical) y los deja colgados, sin adaptarlos ni comentarlos.

Sobre el editor: la responsabilidad la estimo más grave. Nos hallamos ante un libro al que no se le ha dado una forma suficiente; se ha publicado un manuscrito sin escarbar, sin profundizar, sin informarse, sin exigir al autor una confrontación con un músico o un experto sobre la materia. Ha buscado sencillamente la comercialidad del tema, pero no se ha preocupado de su calidad, siquiera científica, como corresponde a una editorial que de tal se precia, y además concede a la obra un premio de importancia nacional en el mundo de la electrónica. Creo que es una gran equivocación, y apelo a este hecho para que en el futuro se produzca mayor calidad, con lo que se benefician todos: el autor, porque será consciente de un nivel, y el editor, por ser un texto comercial. (A.L.R.).

INTRODUCCION A LA GENETICA HUMANA, por A. M. Winchester. Traducción de la segunda edición inglesa por C. Fernández Frías. 216 + XII págs., 75 figuras. Contiene glosario de términos genéticos y tabla alfabética de materias. Editorial Alhambra, 1977. Comienza el doctor Silverio Palafox el prólogo de la versión española afirmando que "Necesitábamos imperiosamente disponer de un libro como éste... porque... son ya cientos... los que precisan una iniciación... a las más diversas cuestiones de la antropobiología –en cuya base figura por derecho propio la genética". Estoy de acuerdo con él en este aserto, ya que, si bien son numerosos los textos de genética humana, sobre todo en inglés, el libro de Winchester viene a llenar un hueco, ofreciendo, en forma simple y asequible, conocimientos avanzados de genética humana. Será de interés este libro para universitarios en general; en particular lo será para quienes cursaron medicina y no recibieron en su carrera la información genética adecuada y que se ven en el momento actual de su práctica bombardeados por términos nuevos, muchos de ellos esotéricos, que han dejado de ser de interés exclusivamente académico, ya que en muchos casos se han hecho problemas de la práctica clínica rutinaria.

En 12 capítulos ha dividido el doctor Winchester su obra: los tres primeros, a modo de introducción, explican el concepto clásico de gen y su localización en los cromosomas, la división celular general (mitosis) y la peculiar (meiosis) que da lugar a los gametos (espermatozoo y óvulo) y, finalmente, el proceso de fecundación y el comienzo de la nueva vida.

Los tres capítulos siguientes forman asimismo un bloque donde se explican, de manera acertadamente simple y comprensible, las leyes fundamentales de la herencia (dominante, recesiva, poligénica y ligada al sexo). Del capítulo séptimo al noveno se pormenorizan tres aspectos clásicos de la genética humana, a saber: los grupos sanguíneos, las enfermedades metabólicas y las variantes de la hemoglobina (elegidas éstas como ejemplo para discutir el fenómeno de las mutaciones).

El capítulo 10 narra sucintamente las alteraciones cromosómicas humanas. El 11 discute brevemente el viejo conflicto herencia-ambiente, con el inevitable tema de los gemelos. Y, finalmente, el capítulo 12 enuncia las principales reglas de la genética de poblaciones.

Al hacer una crítica razonada hemos de ensalzar las virtudes y denunciar las omisiones del manual de Winchester. Empezaré por éstas: *(a)* Creo que el bloque que podríamos llamar de genética clínica (del capítulo siete al nueve) es demasiado reducido. El crecimiento espectacular de nuestros conocimientos en genética de los distintos tipos de padecimientos humanos, bien sean enfermedades congénitas, bien enfermedades comunes del adulto (úlcera, diabetes, infarto, cáncer, etcétera), exige a mi juicio que, incluso una obra titulada "Introducción a la Genética Humana", aborde, siquiera sucintamente, este campo. *(b)* Las alteraciones cromosómicas, reducidas al capítulo 10 y parte del 5, deberían ser tratadas con una extensión mayor, si bien debo reconocer que lo poco que de ellas se dice está bien dicho. *(c)* Finalmente, un libro que busca su público entre el universitario, ávido de conocimientos recientes de un tema tan de actualidad, debería tratar algo más ampliamente el capítulo de genética de poblaciones, si bien, repito lo supradicho, lo poco que se dice está bien dicho.

Virtudes: Fundamentalmente la de lograr que el lenguaje esotérico de la nueva genética se haga asequible a través de la introducción progresiva y razonada de términos y conceptos.

Tengo que decir unas palabras acerca de la traducción. Todo el que haya hecho traducciones sabe cuán difícil es lograr una versión española fiel al estilo del autor y al tiempo legible en castellano, máxime en un terreno donde la mayoría de los términos genéticos no figuran en el diccionario de la lengua castellana. Coincido de nuevo con el doctor Palafox en el elogio de la lograda labor del traductor, con algún fallo aislado, como el de la página 88 ("el príncipe padecía... enfermedad de Bleeder") y con el de ceñirse literalmente al texto en el estilo impersonal, tan corriente en inglés, pero poco elegante en español.

En resumen, el libro de Winchester es una corta pieza de gran valor para los estudiantes universitarios que deseen conocer algo, con poco esfuerzo y en poco tiempo, de lo que es esta nueva y fascinante rama de la genética humana. (A. S. C.).

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

TEORIAS UNIFICADAS DE LAS INTERACCIONES FUNDAMENTALES

HISTORY OF WESTERN PHILOSOPHY. Bertrand Russell.

A SHORT HISTORY OF SOME DISCOVERIES IN ATOMIC PHYSICS. C. N. Yang en *Elementary Particles.*

ELECTRON-POSITRON ANNIHILATION AND THE NEW PARTICLES. S. D. Drell en *Scientific American,* pág. 50; junio, 1975.

EL CONFINAMIENTO DE LOS QUARKS. Y. Nambu en *Investigación y Ciencia,* pág. 52; enero, 1977.

LA CUESTION DEL DIOXIDO DE CARBONO

BROOKHAVEN SYMPOSIA IN BIOLOGY NO. 24: CARBON AND THE BIOSPHERE. Dirigido por George M. Woodwell y E. V. Pecan. U.S. Atomic Energy Commission, Division of Technical Information, Oak Ridge, Tenn., 1973.

PRIMARY PRODUCTION OF INLAND AQUATIC ECOSYSTEMS. Gene E. Likens en *Primary Productivity of the Biosphere,* dirigido por H. Lieth y Robert H. Whittaker. Springer-Verlag, 1975.

LA SUSTITUCION QUIRURGICA DE LA ARTICULACION DE LA RODILLA

POLYCENTRIC KNEE ARTHROPLASTY: PROSTHETIC SIMULATION OF NORMAL KNEE MOVEMENT. Frank H. Gunston en *The Journal of Bone and Joint Surgery,* vol. 53-B. págs. 272-277; mayo, 1971.

GEOMETRIC TOTAL KNEE ARTHROPLASTY, I: CONCEPTION, DESIGN, INDICATIONS, AND SURGICAL TECHNIC. Mark B. Coventry, Jackson E. Upshaw, Lee H. Riley, Gerald A. M. Finerman y Roderick H. Turner en *Clinical Orthopaedics and Related Research,* n.º 94. págs. 171-176; julio-agosto, 1973.

THE SPHEROCENTRIC KNEE: BIOMECHANICAL TESTING AND CLINICAL TRIAL. David A. Sonstegard, Herbert Kaufer y Larry S. Matthews en *The Journal of Bone and Joint Surgery,* vol. 59-A, págs. 602-616; julio, 1977.

LA ESTRUCTURA TRIDIMENSIONAL DEL ARN DE TRANSFERENCIA

THREE-DIMENSIONAL STRUCTURE OF TRANSFER RNA. Sung Hou Kim en *Progress in Nucleic Acid Research and Molecular Biology,* vol. 17, págs. 181-216; 1976.

TRANSFER RNA: MOLECULAR STRUCTURE, SEQUENCE, AND PROPERTIES. Alexander Rich y Uttam L. RajBhandary en *Annual Review of Biochemistry,* vol. 45, págs. 805-860; 1976.

LA ESTRUCTURA DEL MEDIO INTERESTELAR

THE EVOLUTION OF SUPERNOVA REMNANTS, I: SPHERICALLY SYMMETRIC MODELS. Roger A. Chevalier en *The Astrophysical Journal,* vol. 188, n.º 3. parte 1.ª, págs. 501-516; 15 de marzo de 1974.

SOFT X-RAY EMISSION IN ERIDANUS: AN OLD SUPERNOVA REMNANT? S. Naranon, S. Shulman. H. Friedman y G. Fritz en *The Astrophysical Journal,* vol. 208, n.º 3. parte 1.ª, págs. 718-726; 15 de septiembre de 1976.

EL MECANISMO DE ADHERENCIA EN LAS BACTERIAS

STRUCTURE AND FUNCTION OF THE CELL ENVELOPE OF GRAM-NEGATIVE BACTERIA. J. W. Costerton, J. M. Ingram y K.-J. Cheng en *Bacteriological Reviews,* vol. 38. n.º 1, págs. 87-110; marzo, 1974.

DENTAL CARIES. R. J. Gibbons y J. van Houte en *Annual Review of Medicine: Selected Topics in the Clinical Sciences,* vol. 26, págs. 121-136; 1975.

FINE STRUCTURE AND DISTRIBUTION OF EXTRACELLULAR POLYMER SURROUNDING SELECTED AEROBIC BACTERIA. Gerald D. Cagle en *Canadian Journal of Microbiology,* vol. 21, n.º 3, págs. 395-408; marzo, 1975.

LA EFICIENCIA DE LOS ALGORITMOS

ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHEIDUNGSPROBLEM. Alan M. Turing en *Proceedings of the London Mathematical Society,* Series 2, vol. 42, partes 3.ª y 4.ª, págs. 230-265; noviembre/diciembre, 1936.

THE COMPLEXITY OF THEOREM-PROVING PROCEDURES. Stephen A. Cook en *Proceedings of the Third Annual AMC Symposium on Theory of Computing,* págs. 151-178; 1971.

REDUCIBILITY AMONG COMBINATORIAL PROBLEMS. Richard M. Karp en *Complexity of Computer Computations,* dirigido por Raymond E. Miller y James W. Thatcher. Plenum Press, 1972.

TOWARDS FEASIBLE SOLUTIONS OF THE TAUTOLOGY PROBLEM. Bradford Dunham y Hao Wang en *Annals of Mathematical Logic,* vol. 10, n.º 2, págs. 117-154; diciembre, 1976.

LA CARTAGO ROMANA

EXCAVATIONS AT CARTHAGE CONDUCTED BY THE UNIVERSITY OF MICHIGAN, 1975: I. Dirigido por John H. Humphrey. Tunisian National Institute of Archaeology and Art, 1976.

CARTHAGE 1977: THE PUNIC AND ROMAN HARBORS. L. Stager en *Archaeology,* vol. 30, págs. 198-200; 1977.

EXCAVATIONS AT CARTHAGE, 1975: SECOND INTERIM REPORT. Henry Hurst en *Antiquaries Journal,* vol. 56, parte 2.ª, págs. 177-197; 1977.

JUEGOS MATEMATICOS

100 PUZZLES: HOW TO MAKE AND SOLVE THEM. Anthony S. Filipiak. A. S. Barnes and Co., Inc., 1942.

LA SCULPTURE DE BERROCAL. Giuseppe Marchiori. La Connaissance Bruxelles, 1973.

TALLER Y LABORATORIO

A SIMPLE DEVICE FOR OBSERVING THE POLARIZATION OF LIGHT FROM THE SKY. Ulf Halldén en *Nature,* vol. 182. n.º 4631, pág. 333; 2 de agosto de 1958.

HOW BEES NAVIGATE. Allan W. Snyder y Colin Pask en *Nature,* vol. 239, n.º 5366, págs. 48-50; 1 de septiembre de 1972.

ESTIMATING THE SIZE RANGE OF SMOG AEROSOL PARTICLES WITH A PAIR OF SUNGLASSES. W. H. White en *Atmospheric Environment,* vol. 9, n.º 11, págs. 1036-1037; 1975.